moving from start (N-terminus or 5′) to end (C-terminus or 3′), such that for an alignment that extends to p monomers (where $p>x$) there are $p-x+1$ such windows, each window has at least $x \cdot y$ identical aligned monomers, where: x is slected from 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200; y is selected from 0.50, 0.60, 0.70, 0.75, 0.80, 0.85, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99; and if $x \cdot y$ is not an integer then it is rounded up to the nearest integer. The preferred pairwise alignment algorithm is the Needleman-Wunsch global alignment algorithm [Needlman &Wunsch (1970) *J. Mol. Biol.* 48, 443-453], using default parameters (*e.g.*, with Gap opening penalty = 10.0, and with Gap extension penalty = 0.5, using thè EBLOSUM62 scoring matrix). This algorithm is conveniently implemented in the *needle* tool in the EMBOSS package [Rice *et al.* (2000) *Trends Genet.* 16:276-277].

The nucleic acids and polypeptides of the inention may additionally have further sequences to the N-terminus/5′ and/or C-terminus/3′ of these sequences (a) to (d).

All of the Gram positive bacterial sequences referenced herein are publicly available through PubMed on GenBank.

Streptococcus pneumoniae Adhesin Island Sequences

As discussed above, a *S. pneumoniae* AI sequence is present in the TIGR4 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences are set forth below.

SrtD (Sp0468) is a sortase. An example of an amino acid sequence of SrtD is set forth in SEQ ID NO: 80.

SEQ ID NO: 80
MSRTKLRALLGYLLMLVACLIPIYCFGQMVLQSLGQVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

SrtC (Sp0467) is a sortase. An example of an amino acid sequence of SrtC is set forth in SEQ ID NO: 81.

SEQ ID NO: 81
MSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAFNATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIG
YVEIPAIDQEIPMYVGTSEDILQKGAGLLEGASLPVGGENTHTVITAHRGLPTAELFSQLDKMKKGDIFYLHVLD
QVLAYQVDQIVTVEPNDFEPVLIQHGEDYATLLTCTPYMINSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWL
LLGAMAVILLLLLYRVYRNRRIVKGLEKQLEGRHVKD

SrtB (SP0466) is a sortase. An example of an amino acid sequence of SrtB is set forth in SEQ ID NO: 82.

SEQ ID NO: 82
MAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDSLNNVVSGDPWSEEMKKKGRAEYARM
LEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTHAVITAHTGLPTAKMFTDLTKLKVGD
KFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINTHRLLVRGHRIPYVAEVEEEFIAANK
LSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVEDGQQ

Sp0465 is a hypothetical protein. An example of an amino acid sequence of Sp0465 is set forth in SEQ ID NO: 83.

SEQ ID NO: 83
MFLPFLSASLYLQTHHFIAFPNRQSYLLRETRKSHFFLIHHPF

RrgC (SP0464) is a cell wall surface anchor family protein. RrgC contains a sortase substrate motif VPXTG (SEQ ID NO: 137), shown in italics in SEQ ID NO: 84.

SEQ ID NO: 84
MISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRVQIVRDLHS
WDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAKKTDTMTTK
VKLIKVDQDHNRLEGVGFKLVSVARDVSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGNYRFKEVEP
LAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVLQNGKEVVV
TSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRID*VPDTG*EETLYILML
VAILLFGSGYYLTKKPNN

RrgB (Sp0463) is a cell wall surface anchor protein. RrgB contains a sortase substrate motif IPXTG (SEQ ID NO: 133), shown in italics in SEQ ID NO: 85.

SEQ ID NO: 85
MKSINKFLTMLAALLLTASSLFSAATVFAAGTTTTSVTVHKLLATDGDMDKIANELETGNYAGNKVGVLPANAKE
IAGVMFVWTNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAAKYKIYEIHSLSTY
VGEDGATLTGSKAVPIEIELPLNDVVDAHVYPKNTEAKPKIDKDFKGKANPDTPRVDKDTPVNHQVGDVVEYEIV
TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDDVALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKTVKI
TYSATLNDKAIVEVPESNDVTFNYGNNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPIPAGAEATFDLVNAQTG
KVVQTVTLTTDKNTVTVNGLDKNTEYKFVERSIKGYSADYQEITTAGEIAVKNWKDENPKPLDPTEPKVVTYGKK
FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
AQAAYNAAVIAANNAFEWVADKDNENVVKLVSDAQGRFEITGLLAGTYYLEETKQPAGYALLTSRQKFEVTATSY
SATGQGIEYTAGSGKDDATKVVNKKIT*IPQTG*GIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

RrgA (Sp0462) is a cell wall surface anchor protein. RrgA contains a sortase substrate motif YPXTG (SEQ ID NO: 186), indicated in italics in SEQ ID NO: 86.

SEQ ID NO: 86
MLNRETHMKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDG
TTVSQRTEAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGT
YPDVQTPYQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTVYEQKDKSVPL
DVVILLDNSNSMSNIRNKNARRAERAGEATRSLIDKITSDSENRVALVTYASTIFDGTEFTVEKGVADKNGKRLN
DSLFWNYDQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQQAR
QNSQKVIFHITDGVPTMSYPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDGQSYQM
FTDKTVYEKGAPAAFPVKPEKYSEMKAAGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGDPTRWYYNG
NIAPDGYDVFTVGIGINGDPGTDEATATSFMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENGTITD
PMGELIDLQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNAKVLYDTTEKRIRVTGLYLGTDEKVT
LTYNVRLNDEFVSNKFYDTNGRTTLHPKEVEQNTVRDFPIPKIRDVRKYPEITISKEKKLGDIEFIKVNKNDKKP
LRGAVFSLQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVN
GEVRDVTSIVPQDIPAGYEFTNDKHYITNEPIPPKRE*YPRTG*GIGMLPFYLIGCMMMGGVLLYTRKHP

RlrA (Sp0461) is a transcriptional regulator. An example of an amino acid sequence for RlrA is set forth in SEQ ID NO: 87.

SEQ ID NO: 87
MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

As discussed above, a *S. pneumoniae* AI sequence is present in the *S. pneumoniae* strain 670 genome. Examples of *S. pneumoniae* AI sequences are set forth below.

Orf1_670 is a transposase. An example of an amino acid sequence of orf1_670 is set forth in SEQ ID NO: 171.

SEQ ID NO: 171

MEHINHTTLLIGIKDKNITLNKAIQHDTHIEVFATLDYHPPKCKHCKGKQIKYDFQKPSKIPFIEIGGFPSLIHL
KKRRFQCKSCRKVTVAETTLVQKNCQISEMVRQKIAQLLLNREALTHIASKLAISTSTSTVYRKLKQFHFQEDYT
TLPEILSWDEFSYQKGKLAFIAQDFNTKKIMTILDNRRQTTIRNHFFKYSKEARKKVKVVTVDMSGSYIPLIKKL
FPNAKIVLDRFHIVQHMSRALNQTRINIMKQFDDKSLEYRALKYYWKFILKDSRKLSLKPFYARTFRETLTPREC
LKKIFTLVPELKDYYDLYQLLLFHLQEKNTDQFWGLIQDTLPHLNRTFKTTLSTFICYKNYITNAIELPYSNAKL
EATNKLIKDIKRNAFGFRNFENFKKRIFIALNIKKERTKFVLSRA

Orf2_670 is a transcriptional regulator. An example of an amino acid sequence of Orf2_670 is set forth in SEQ ID NO: 172.

SEQ ID NO: 172

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

Orf3_670 is a cell wall surface anchor family proten. An example of an amino acid sequence of Orf3_670 is set forth in SEQ ID NO: 173.

SEQ ID NO: 173

MLNRETHMKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDG
TTVSQRTEAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGT
YPDVQTPYQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPL
DVVILLDNSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILN
DSALWTFDRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQAR
PNSKKVIFHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQM
FTKKPVTDQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMA
QDGYDVFTVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMG
ELIDFQLGADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTY
NVRLNDQFVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRD
AVFSLQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEV
RDVTSIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

Orf4_670 is a cell wall surface anchor family protein. An example of an amino acid sequence of orf4_670 is set forth in SEQ ID NO: 174.

SEQ ID NO: 174

MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGKLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

Orf5_670 is a cell wall surface anchor family protein. An example of an amino acid sequence of orf5_670 is set forth in SEQ ID NO: 175.

SEQ ID NO: 175

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK

KTDTMTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

Orf6_670 is a sortase. An example of an amino acid sequence of orf6_670 is set forth in SEQ ID NO: 176.

SEQ ID NO: 176

MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

Orf7_670 is a sortase. An example of an amino acid sequence of orf7_670 is set forth in SEQ ID NO: 177.

SEQ ID NO: 177

VSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAFNATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIG
YVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGENTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLD
QVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYMINSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWL
LLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

Orf8_670 is a sortase. An example of an amino acid sequence of orf8_670 is set forth in SEQ ID NO: 178.

SEQ ID NO: 178

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 19A Hungary 6 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 19A Hungary 6 are set forth below.

ORF2_19AH is a transcriptional regulator. An example of an amino acid sequence of ORF2_19AH is set forth in SEQ ID NO: 187.

SEQ ID NO: 187

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_19AH is a cell wall surface protein. An example of an amino acid sequence of ORF3_19AH is set forth in SEQ ID NO: 188.

SEQ ID NO: 188

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
NSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT

DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTSIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKNP

ORF4_19AH is a cell wall surface protein. An example of an amino acid sequence of ORF4_19AH is set forth in SEQ ID NO: 189.

SEQ ID NO: 189
MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGXNGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGKLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

ORF5_19AH is a cell wall surface protein. An example of an amino acid sequence of ORF5_19AH is set forth in SEQ ID NO: 190.

SEQ ID NO: 190
MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KTDTMTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

ORF6_19AH is a putative sortase. An example of an amino acid sequence of ORF6_19AH is set forth in SEQ ID NO: 191.

SEQ ID NO: 191
MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

ORF7_19AH is a putative sortase. An example of an amino acid sequence of ORF7_19AH is set forth in SEQ ID NO: 192.

SEQ ID NO: 192
MDNSRRSRKKGTKKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

ORF8_19AH is a putative sortase. An example of an amino acid sequence of ORF8_19AH is set forth in SEQ ID NO: 193.

SEQ ID NO: 193
MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFMGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 6B Finland 12 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 6B Finland 12 are set forth below.

ORF2_6BF is a transcriptional regulator. An example of an amino acid sequence of ORF2_6BF is set forth in SEQ ID NO: 194.

SEQ ID NO: 194

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_6BF is a cell wall surface protein. An example of an amino acid sequence of ORF3_6BF is set forth in SEQ ID NO: 195.

SEQ ID NO: 195

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
NSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTSIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

ORF4_6BF is a cell wall surface protein. An example of an amino acid sequence of ORF4_6BF is set forth in SEQ ID NO: 196.

SEQ ID NO: 196

MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGKLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

ORF5_6BF is a cell wall surface protein. An example of an amino acid sequence of ORF5_6BF is set forth in SEQ ID NO: 197.

SEQ ID NO: 197

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KTDTMTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

ORF6_6BF is a putative sortase. An example of an amino acid sequence of ORF6_6BF is set forth in SEQ ID NO: 198.

SEQ ID NO: 198

MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

ORF7_6BF is a putative sortase. An example of an amino acid sequence of ORF7_6BF is set forth in SEQ ID NO: 199.

SEQ ID NO: 199

MDNSRRSRKKGTKKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

ORF8_6BF is a putative sortase. An example of an amino acid sequence of ORF8_6BF is set forth in SEQ ID NO: 200.

SEQ ID NO: 200

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 6B Spain 2 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 6B Spain 2 are set forth below.

ORF2_6BSP is a transcriptional regulator. An example of an amino acid sequence of ORF2_6BSP is set forth in SEQ ID NO: 201.

SEQ ID NO: 201

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_6BSP is a cell wall surface protein. An example of an amino acid sequence of ORF3_6BSP is set forth in SEQ ID NO: 202.

SEQ ID NO: 202

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
NSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTSIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

ORF4_6BSP is a cell wall surface protein. An example of an amino acid sequence of ORF4_6BSP is set forth in SEQ ID NO: 203.

SEQ ID NO: 203

MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGKLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

ORF5_6BSP is a cell wall surface protein. An example of an amino acid sequence of ORF5_6BSP is set forth in SEQ ID NO: 204.

SEQ ID NO: 204

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KTDTMTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

ORF6_6BSP is a putative sortase. An example of an amino acid sequence of ORF6_6BSP is set forth in SEQ ID NO: 205.

SEQ ID NO: 205

MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

ORF7_6BSP is a putative sortase. An example of an amino acid sequence of ORF7_6BSP is set forth in SEQ ID NO: 206.

SEQ ID NO: 206

MDNSRRSRKKGTKKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

ORF8_6BSP is a putative sortase. An example of an amino acid sequence of ORF8_6BSP is set forth in SEQ ID NO: 207.

SEQ ID NO: 207

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 9V Spain 3 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 9V Spain 3 are set forth below.

ORF2_9VSP is a transcriptional regulator. An example of an amino acid sequence of ORF2_9VSP is set forth in SEQ ID NO: 208.

SEQ ID NO: 208
MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_9VSP is a cell wall surface protein. An example of an amino acid sequence of ORF3_9VSP is set forth in SEQ ID NO: 209.

SEQ ID NO: 209
MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTNGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQRTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTVYERKDKSVPLDVVILLD
NSNSMSNIRNKNARRAERAGEATRSLIDKITSDPENRVALVTYASTIFDGTEFTVEKGVADKNGKRLNDSLFWNY
DQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQQARQNSQKVI
FHITDGVPTMSYPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDGQSYQMFTDKTVY
EKGAPAAFPVKPEKYSEMKAVGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGDPTRWYYNGNIAPDGY
DVFTVGIGINGDPGTDEATATSFMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENGTITDPMGELID
LQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRL
NDQFVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPAITIAKEKKLGEIEFIKINKNDKKPLRDAVFS
LQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
SIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLLFYLIGCMMMGGVLLYTRKHP

ORF4_9VSP is a cell wall surface protein. An example of an amino acid sequence of ORF4_9VSP is set forth in SEQ ID NO: 210.

SEQ ID NO: 210
MKSINKFLTMLAALLLTASSLFSAATVFAAGTTTTSVTVHKLLATDGDMDKIANELETGNYAGNKVGVLPANAKE
IAGVMFVWTNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAAKYKIYEIHSLSTY
VGEDGATLTGSKAVPIEIELPLNDVVDAHVYPKNTEAKPKIDKDFKGKANPDTPRVDKDTPVNHQVGDVVEYEIV
TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDDVALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKTVKI
TYSATLNDKAIVEVPESNDVTFNYGNNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPIPAGAEATFDLVNAQTG
KVVQTVTLTTDKNTVTVNGLDKNTEYKFVERSIKGYSADYQEITTAGEIAVKNWKDENPKPLDPTEPKVVTYGKK
FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
AQAAYNAAVIAANNAFEWVADKDNENVVKLVSDAQGRFEITGLLAGTYYLEETKQPAGYALLTSRQKFEVTATSY
SATGQGIEYTAGSGKDDATKVVNKKITIPQTGGIGTIIFAVAGAVIMGIAVYAYVKNNKDEDQLA

ORF5_9VSP is a cell wall surface protein. An example of an amino acid sequence of ORF5_9VSP is set forth in SEQ ID NO: 211.

SEQ ID NO: 211
MTMQKMQKMQKMQKMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVW
KLDDSYSYDNRVQIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMT
DQTVEPLVIVAKKADTVTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKN
GEIVVTNLPLGTYRFKEVEPLAGYTVTTMDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKV
MKEENGHYTPVLQNGKEVVVASGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNN
KRPRIDVPDTGEETLYILMLVAILLFGSGYYLTKKTNN

ORF6_9VSP is a putative sortase. An example of an amino acid sequence of ORF6_9VSP is set forth in SEQ ID NO: 212.

SEQ ID NO: 212
MLIKMAKTKKQKRNNLLLGVVFFIGIAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPAIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKRQSERALKALKEATKEVKVE
DE

ORF7_9VSP is a putative sortase. An example of an amino acid sequence of ORF7_9VSP is set forth in SEQ ID NO: 213.

SEQ ID NO: 213
MSKSRYSRKKSVKKKKNPFILLLIFLVGLAVAMYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDIFYLHVLDQVLAYQVDQIVTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLGAMAVILLLLYRVYRNRRIVKGLEKQLEGRHVKD

ORF8_9VSP is a putative sortase. An example of an amino acid sequence of ORF8_9VSP is set forth in SEQ ID NO: 214.

SEQ ID NO: 214
MSRTKLRALLGYLLMLVACLIPIYCFGQMVLQSLGQVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 14 CSR 10 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 14 CSR 10 are set forth below.

ORF2_14CSR is a transcriptional regulator. An example of an amino acid sequence of ORF2_14CSR is set forth in SEQ ID NO: 215.

SEQ ID NO: 215
MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_14CSR is a cell wall surface protein. An example of an amino acid sequence of ORF3_14CSR is set forth in SEQ ID NO: 216.

SEQ ID NO: 216
MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
NSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTSIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

ORF4_14CSR is a cell wall surface protein. An example of an amino acid sequence of ORF4_14CSR is set forth in SEQ ID NO: 217.

SEQ ID NO: 217
MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGNNGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE

NNWTYTWSGLDNSTEYKVEEHYNGYSAEYTVESKGKLGVKNWKDNNPAPINPEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAAIMGIAVYAYVKNNKDEDQLA

ORF5_14CSR is a cell wall surface protein. An example of an amino acid sequence of ORF5_14CSR is set forth in SEQ ID NO: 218.

SEQ ID NO: 218
MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDDRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KTDTMTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIFVTNLPLGN
YRFKEVEPLAGYAVTTLDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEESGHYTPVL
QNGKEVVVTSGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKPNN

ORF6_14CSR is a putative sortase. An example of an amino acid sequence of ORF6_14CSR is set forth in SEQ ID NO: 219.

SEQ ID NO: 219
MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

ORF7_14CSR is a putative sortase. An example of an amino acid sequence of ORF7_14CSR is set forth in SEQ ID NO: 220.

SEQ ID NO: 220
MDNSRRSRKKGTKKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

ORF8_14CSR is a putative sortase. An example of an amino acid sequence of ORF8_14CSR is set forth in SEQ ID NO: 221.

SEQ ID NO: 221
MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 19F Taiwan 14 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 19F Taiwan 14 are set forth below.

ORF2_19FTW is a transcriptional regulator. An example of an amino acid sequence of ORF2_19FTW is set forth in SEQ ID NO: 222.

SEQ ID NO: 222
MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF3_19FTW is set forth in SEQ ID NO: 223.

SEQ ID NO: 223

MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTVYERKDKSVPLDVVILLD
NSNSMSNIRNKNARRAERAGEATRSLIDKITSDPENRVALVTYASTIFDGTEFTVEKGVADKNGKRLNDSLFWNY
DQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQQARQNSQKVI
FHITDGVPTMSYPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDGQSYQMFTDKTVY
EKGAPAAFPVKPEKYSEMKAVGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGAPTRWYYNGNIAPDGY
DVFTVGIGINGDPGTDEATATSFMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENGTITDPMGELID
LQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRL
NDQFVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPAITIAKEKKLGEIEFIKINKNDKKPLRDAVFS
LQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
SIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

ORF4_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF4_19FTW is set forth in SEQ ID NO: 224.

SEQ ID NO: 224

MKSINKFLTMLAALLLTASSLFSAATVFAAGTTTTSVTVHKLLATDGDMDKIANELETGNYAGNKVGVLPANAKE
IAGVMFVWTNTNNEIIDENGQTLGVNIDPQTFKLSGAMPATAMKKLTEAEGAKFNTANLPAAKYKIYEIHSLSTY
VGEDGATLTGSKAVPIEIELPLNDVVDAHVYPKNTEAKPKIDKDFKGKANPDTPRVDKDTPVNHQVGDVVEYEIV
TKIPALANYATANWSDRMTEGLAFNKGTVKVTVDDVALEAGDYALTEVATGFDLKLTDAGLAKVNDQNAEKTVKI
TYSATLNDKAIVEVPESNDVTFNYGNNPDHGNTPKPNKPNENGDLTLTKTWVDATGAPIPAGAEATFDLVNAQTG
KVVQTVTLTTDKNTVTVNGLDKNTEYKFVERSIKGYSADYQEITTAGEIAVKNWKDENPKPLDPTEPKVVTYGKK
FVKVNDKDNRLAGAEFVIANADNAGQYLARKADKVSQEEKQLVVTTKDALDRAVAAYNALTAQQQTQQEKEKVDK
AQAAYNAAVIAANNAFEWVADKDNENVVKLVSDAQGRFEITGLLAGTYYLEETKQPAGYALLTSRQKFEVTATSY
SATGQGIEYTAGSGKDDATKVVNKKITIPQTGGIGTIIFAVAGAVIMGIAVYAYVKNNKDEDQLA

ORF5_19FTW is a cell wall surface protein. An example of an amino acid sequence of ORF5_19FTW is set forth in SEQ ID NO: 225.

SEQ ID NO: 225

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDNRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KADTVTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIVVTNLPLGT
YRFKEVEPLAGYTVTTMDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
QNGKEVVVASGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKTNN

ORF6_19FTW is a putative sortase. An example of an amino acid sequence of ORF6_19FTW is set forth in SEQ ID NO: 226.

SEQ ID NO: 226

MLIKMAKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPAIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKRQSERALKALKEATKEVKVE
DE

ORF7_19FTW is a putative sortase. An example of an amino acid sequence of ORF7_19FTW is set forth in SEQ ID NO: 227.

SEQ ID NO: 227

MSKSRYSRKKSVKKKKNPFILLLIFLVGLAVAMYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTDQEKKQGVSEYANMLKVHERIGYVEIPAIEQEIPMYVGTSEDILQKGAGLLEGASLPVGGE

NTHLVITAHRGLPTAELFSQLDKMKKGDIFYLHVLDQVLAYQVDQIVTVEPNDFEPVLIQHGQDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLGAMAVILLLLYRVYRNRRIVKGLEKQLEGRHVKD

ORF8_19FTW is a putative sortase. An example of an amino acid sequence of ORF8_19FTW is set forth in SEQ ID NO: 228.

SEQ ID NO: 228
MSRTKLRALLGYLLMLVACLIPIYCFGQMVLQSLGQVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 23F Taiwan 15 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 23F Taiwan 15 are set forth below.

ORF2_23FTW is a transcriptional regulator. An example of an amino acid sequence of ORF2_23FTW is set forth in SEQ ID NO: 229.

SEQ ID NO: 229
MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_23FTW is a cell wall surface protein. An example of an amino acid sequence of ORF3_23FTW is set forth in SEQ ID NO: 230.

SEQ ID NO: 230
MKKVRKIFQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTVYEQKDKSVPLDVVILLD
NSNSMSNIRNKNARRAERAGEATRSLIDKITSDPENRVALVTYASTIFDGTEFTVEKGVADKNGKRLNDSLFWNY
DQTSFTTNTKDYSYLKLTNDKNDIVELKNKVPTEAEDHDGNRLMYQFGATFTQKALMKADEILTQQARQNSQKVI
FHITDGVPTMSYPINFNHATFAPSYQNQLNAFFSKSPNKDGILLSDFITQATSGEHTIVRGDGQSYQMFTDKTVY
EKGAPAAFPVKPEKYSEMKAAGYAVIGDPINGGYIWLNWRESILAYPFNSNTAKITNHGDPTRWYYNGNIAPDGY
DVFTVGIGINGDPGTDEATATSFMQSISSKPENYTNVTDTTKILEQLNRYFHTIVTEKKSIENGTITDPMGELID
LQLGTDGRFDPADYTLTANDGSRLENGQAVGGPQNDGGLLKNAKVLYDTTEKRIRVTGLYLGTDEKVTLTYNVRL
NDEFVSNKFYDTNGRTTLHPKEVEQNTVRDFPIPKIRDVRKYPEITISKEKKLGDIEFIKVNKNDKKPLRDAVFS
LQKQHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVT
SIVPQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKHP

ORF4_23FTW is a cell wall surface protein. An example of an amino acid sequence of ORF4_23FTW is set forth in SEQ ID NO: 231.

SEQ ID NO: 231
MKSINKFLTILAALLLTVSSLFSAATVFAAEQKTKTLTVHKLLMTDQELDAWNSDAITTAGYDGSQNFEQFKQLQ
GVPQGVTEISGVAFELQSYTGPQGKEQENLTNDAVWTAVNKGVTTETGVKFDTEVLQGTYRLVEVRKESTYVGPN
GKVLTGMKAVPALITLPLVNQNGVVENAHVYPKNSEDKPTATKTFDTAAGFVDPGEKGLAIGTKVPYIVTTTIPK
NSTLATAFWSDEMTEGLDYNGDVVVNYNGQPLDNSHYTLEAGHNGFILKLNEKGLEAINGKDAEATITLKYTATL
NALAVADVPEANDVTFHYGNNPGHGNTPKPNKPKNGELTITKTWADAKDAPIAGVEVTFDLVNAQTGEVVKVPGH
ETGIVLNQTNNWTFTATGLDNNTEYKFVERTIKGYSADYQTITETGKIAVKNWKDENPEPINPEEPRVKTYGKKF
VKVDQKDERLKEAQFVVKNEQGKYLALKSAAQQAVNEKAAAEAKQALDAAIAAYTNAADKNAAQAVVDAAQKTYN
DNYRAARFGYVEVERKEDALVLTSNTDGQFQISGLAAGSYTLEETKAPEGFAKLGDVKFEVGAGSWNQGDFNYLK
DVQKNDATKVVNKKITIPQTGGIGTIIFAVAGAVIMGIAVYAYVKNNKDEDQLA

ORF5_23FTW is a cell wall surface protein. An example of an amino acid sequence of ORF5_23FTW is set forth in SEQ ID NO: 232.

SEQ ID NO: 232

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDNRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KADTVTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIVVTNLPLGT
YRFKEVEPLAGYTVTTMDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
QNGKEVVVASGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKTNN

ORF6_23FTW is a putative sortase. An example of an amino acid sequence of ORF6_23FTW is set forth in SEQ ID NO: 233.

SEQ ID NO: 233

MLIKMVKTKKQKRNNLLLGVVFFIGMAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPVIDVDLPVYAGTAEEVLQQGAGQLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKKQPEKALKALKAARKEVKVE
DGQQ

ORF7_23FTW is a putative sortase. An example of an amino acid sequence of ORF7_23FTW is set forth in SEQ ID NO: 234.

SEQ ID NO: 234

MDNSRRSRKKGTKKKKHPLILLLIFLVGFAVAIYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDVFYLHVLDQVLAYQVDQILTVEPNDFEPVLIQHGKDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLAALVMILVLSYGVYRHRRIVKGLEKQLEEHHVKG

ORF8_23FTW is a putative sortase. An example of an amino acid sequence of ORF8_23FTW is set forth in SEQ ID NO: 235.

SEQ ID NO: 235

MSKAKLQKLLGYLLMLVALVIPVYCFGQMVLQSLGQVKGHEIFSESVTADSYQEQLQRSLDYNQRLDSQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLAMGLAHVDGTPLPVEGKGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

As discussed above, a *S. pneumoniae* AI sequence is present in the 23F Poland 16 *S. pneumoniae* genome. Examples of *S. pneumoniae* AI sequences from 23F Poland 16 are set forth below.

ORF2_23FP is a transcriptional regulator. An example of an amino acid sequence of ORF2_23FP is set forth in SEQ ID NO: 236.

SEQ ID NO: 236

MLNKYIEKRITDKITILNILLDIRSIELDELSTLTSLQSKSLLSILQELQETFEEELTFNLDTQQVQLIEHHSHQ
TNYYFHQLYNQSTILKILRFFLLQGNQSFNEFTQKEYISIATGYRVRQKCGLLLRSVGLDLVKNQVVGPEYRIRF
LIALLQFHFGIEIYDLNDGSMDWVTHMIVQSNSQLSHELLEITPDEYVHFSILVALTWKRREFPLEFPESKEFEK
LKNLFMYPILMEHCQTYLEPHANMTFTQEELDYIFLVYCSANSSFSKDKWNQEKKTHTIQLILQHTRGKHLLSKF
KNILGNDISNSLSFLTALTFLTRTFLFGLQNLVPYYNYYEHYGIESDKPLYHISKAIVQEWMTEQKIEGVIDQHR
LYLFSLYLTETIFSSLPAIPIFIILNNQADVNLIKSIILRNFTDKVASVTGYNILISPPPSEEHLTEPLIIITTK
EYLPYVKKQYPKGKHHFLTIALDLHVSQQRLIYQTIVDIRKEAFDKRVAMIAKKAHYLL

ORF3_23FP is a cell wall surface protein. An example of an amino acid sequence of ORF3_23FP is set forth in SEQ ID NO: 237.

SEQ ID NO: 237

MKKVRKIEQKAVAGLCCISQLTAFSSIVALAETPETSPAIGKVVIKETGEGGALLGDAVFELKNNTDGTTVSQRT
EAQTGEAIFSNIKPGTYTLTEAQPPVGYKPSTKQWTVEVEKNGRTTVQGEQVENREEALSDQYPQTGTYPDVQTP
YQIIKVDGSEKNGQHKALNPNPYERVIPEGTLSKRIYQVNNLDDNQYGIELTVSGKTTVETKEASTPLDVVILLD
NSNSMSNIRHNHAHRAEKAGEATRALVDKITSNPDNRVALVTYGSTIFDGSEATVEKGVADANGKILNDSALWTF
DRTTFTAKTYNYSFLNLTSDPTDIQTIKDRIPSDAEELNKDKLMYQFGATFTQKALMTADDILTKQARPNSKKVI
FHITDGVPTMSYPINFKYTGTTQSYRTQLNNFKAKTPNSSGILLEDFVTWSADGEHKIVRGDGESYQMFTKKPVT
DQYGVHQILSITSMEQRAKLVSAGYRFYGTDLYLYWRDSILAYPFNSSTDWITNHGDPTTWYYNGNMAQDGYDVF
TVGVGVNGDPGTDEATATRFMQSISSSPDNYTNVADPSQILQELNRYFYTIVNEKKSIENGTITDPMGELIDFQL
GADGRFDPADYTLTANDGSSLVNNVPTGGPQNDGGLLKNAKVFYDTTEKRIRVTGLYLGTGEKVTLTYNVRLNDQ
FVSNKFYDTNGRTTLHPKEVEKNTVRDFPIPKIRDVRKYPEITIPKEKKLGEIEFIKINKNDKKPLRDAVFSLQK
QHPDYPDIYGAIDQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPIVAFQIVNGEVRDVTSIV
PQDIPAGYEFTNDKHYITNEPIPPKREYPRTGGIGMLPFYLIGCMMMGGVLLYTRKNP

ORF4_23FP is a cell wall surface protein. An example of an amino acid sequence of ORF4_23FP is set forth in SEQ ID NO: 238.

SEQ ID NO: 238

MKSINKFLTMLAALLLTASSLFSAATVFAADNVSTAPDAVTKTLTIHKLLLSEDDLKTWDTNGPKGYDGTQSSLK
DLTGVVAEEIPNVYFELQKYNLTDGKEKENLKDDSKWTTVHGGLTTKDGLKIETSTLKGVYRIREDRTKTTYVGP
NGQVLTGSKAVPALVTLPLVNNNGTVIDAHVFPKNSYNKPVVDKRIADTLNYNDQNGLSIGTKIPYVVNTTIPSN
ATFATSFWSDEMTEGLTYNEDVTITLNNVAMDQADYEVTKGINGFNLKLTEAGLAKINGKDADQKIQITYSATLN
SLAVADIPESNDITYHYGNHQDHGNTPKPTKPNNGQITVTKTWDSQPAPEGVKATVQLVNAKTGEKVGAPVELSE
NNWTYTWSGLDNSIEYKVEEEYNGYSAEYTVESKGKLGVKNWKDNNPAPINLEEPRVKTYGKKFVKVDQKDTRLE
NAQFVVKKADSNKYIAFKSTAQQAADEKAAATAKQKLDAAVAAYTNAADKQAAQALVDQAQQEYNVAYKEAKFGY
VEVAGKDEAMVLTSNTDGQFQISGLAAGTYKLEEIKAPEGFAKIDDVEFVVGAGSWNQGEFNYLKDVQKNDATKV
VNKKITIPQTGGIGTIIFAVAGAVIMGIAVYAYVKNNKDEDQLA

ORF5_23FP is a cell wall surface protein. An example of an amino acid sequence of ORF5_23FP is set forth in SEQ ID NO: 239.

SEQ ID NO: 239

MTMQKMQKMISRIFFVMALCFSLVWGAHAVQAQEDHTLVLQLENYQEVVSQLPSRDGHRLQVWKLDDSYSYDNRV
QIVRDLHSWDENKLSSFKKTSFEMTFLENQIEVSHIPNGLYYVRSIIQTDAVSYPAEFLFEMTDQTVEPLVIVAK
KADTVTTKVKLIKVDQDHNRLEGVGFKLVSVARDGSEKEVPLIGEYRYSSSGQVGRTLYTDKNGEIVVTNLPLGT
YRFKEVEPLAGYAVTTMDTDVQLVDHQLVTITVVNQKLPRGNVDFMKVDGRTNTSLQGAMFKVMKEENGHYTPVL
QNGKEVVVASGKDGRFRVEGLEYGTYYLWELQAPTGYVQLTSPVSFTIGKDTRKELVTVVKNNKRPRIDVPDTGE
ETLYILMLVAILLFGSGYYLTKKTNN

ORF6_23FP is a putative sortase. An example of an amino acid sequence of ORF6_23FP is set forth in SEQ ID NO: 240.

SEQ ID NO: 240

MLIKMAKTKKQKRNNLLLGVVFFIGIAVMAYPLVSRLYYRVESNQQIADFDKEKATLDEADIDERMKLAQAFNDS
LNNVVSGDPWSEEMKKKGRAEYARMLEIHERMGHVEIPAIDVDLPVYAGTAEEVLQQGAGHLEGTSLPIGGNSTH
AVITAHTGLPTAKMFTDLTKLKVGDKFYVHNIKEVMAYQVDQVKVIEPTNFDDLLIVPGHDYVTLLTCTPYMINT
HRLLVRGHRIPYVAEVEEEFIAANKLSHLYRYLFYVAVGLIVILLWIIRRLRKKKRQSERALKALKEATKEVKVE
DE

ORF7_23FP is a putative sortase. An example of an amino acid sequence of ORF7_23FP is set forth in SEQ ID NO: 241.

SEQ ID NO: 241

MSKSRYSRKKSVKKKKNPFILLLIFLVGLAVAMYPLVSRYYYRIESNEVIKEFDETVSQMDKAELEERWRLAQAF
NATLKPSEILDPFTEQEKKKGVSEYANMLKVHERIGYVEIPAIDQEIPMYVGTSEEILQKGAGLLEGASLPVGGE
NTHTVVTAHRGLPTAELFSQLDKMKKGDIFYLHVLDQVLAYQVDQIVTVEPNDFEPVLIQHGEDYATLLTCTPYM
INSHRLLVRGKRIPYTAPIAERNRAVRERGQFWLWLLLGAMAVILLLLYRVYRNRRIVKGLEKQLEGRHVKD

ORF8_23FP is a putative sortase. An example of an amino acid sequence of ORF8_23FP is set forth in SEQ ID NO: 242.

SEQ ID NO: 242

MSRTKLRALLGYLLMLVACLIPIYCFGQMVLQSLGQVKGHATFVKSMTTEMYQEQQNHSLAYNQRLASQNRIVDP
FLAEGYEVNYQVSDDPDAVYGYLSIPSLEIMEPVYLGADYHHLGMGLAHVDGTPLPLDGTGIRSVIAGHRAEPSH
VFFRHLDQLKVGDALYYDNGQEIVEYQMMDTEIILPSEWEKLESVSSKNIMTLITCDPIPTFNKRLLVNFERVAV
YQKSDPQTAAVARVAFTKEGQSVSRVATSQWLYRGLVVLAFLGILFVLWKLARLLRGK

Immunogenic compositions of the invention comprising AI antigens may further comprise one or more antigenic agents. Preferred antigens include those listed below. Additionally, the compositions of the present invention may be used to treat or prevent infections caused by any of the below-listed microbes. Antigens for use in the immunogenic compositions include, but are not limited to, one or more of the following set forth below, or antigens derived from one or more of the following set forth below:

Bacterial Antigens

N. meningitides: a protein antigen from *N. meningitides* serogroup A, C, W135, Y, and/or B (1-7); an outer-membrane vesicle (OMV) preparation from *N. meningitides* serogroup B. (8, 9, 10, 11); a saccharide antigen, including LPS, from *N. meningitides* serogroup A, B, C W135 and/or Y, such as the oligosaccharide from serogroup C (see PCT/US99/09346; PCT IB98/01665; and PCT IB99/00103);

Streptococcus pneumoniae: a saccharide or protein antigen, particularly a saccharide from *Streptooccus pneumoniae*;

Streptococcus agalactiae: particularly, Group B streptococcus antigens;

Streptococcus pyogenes: particularly, Group A streptococcus antigens;

Enterococcus faecalis or Enterococcus faecium: Particularly a trisaccharide repeat or other *Enterococcus* derived antigens provided in US Patent No. 6,756,361;

Helicobacter pylori: including: Cag, Vac, Nap, HopX, HopY and/or urease antigen;

Bordetella pertussis: such as petussis holotoxin (PT) and filamentous haemagglutinin (FHA) from *B. pertussis*, optionally also combination with pertactin and/or agglutinogens 2 and 3 antigen;

Staphylococcus aureus: including *S. aureus* type 5 and 8 capsular polysaccharides optionally conjugated to nontoxic recombinant *Pseudomonas aeruginosa* exotoxin A, such as StaphVAX™, or antigens derived from surface proteins, invasins (leukocidin, kinases, hyaluronidase), surface factors that inhibit phagocytic engulfment (capsule, Protein A), carotenoids, catalase production, Protein A, coagulase, clotting factor, and/or membrane-damaging toxins (optionally detoxified) that lyse eukaryotic cell membranes (hemolysins, leukotoxin, leukocidin);

Staphylococcus epidermis: particularly, *S. epidermidis* slime-associated antigen (SAA);

Staphylococcus saprophyticus: (causing urinary tract infections) particularly the 160 kDa hemagglutinin of *S. saprophyticus* antigen;

Pseudomonas aeruginosa: particularly, endotoxin A, Wzz protein, *P. aeruginosa* LPS, more particularly LPS isolated from PAO1 (O5 serotype), and/or Outer Membrane Proteins, including Outer Membrane Proteins F (OprF) *(Infect Immun.* 2001 May; 69(5): 3510-3515);

Bacillus anthracis (anthrax): such as *B. anthracis* antigens (optionally detoxified) from A-components (lethal factor (LF) and edema factor (EF)), both of which can share a common B-component known as protective antigen (PA);

Moraxella catarrhalis: (respiratory) including outer membrane protein antigens (HMW-OMP), C-antigen, and/or LPS;

Yersinia pestis (plague): such as F1 capsular antigen (*Infect Immun.* 2003 Jan; 71(1)): 374-383, LPS (*Infect Immun.* 1999 Oct; 67(10): 5395), Yersinia pestis V antigen (*Infect Immun.* 1997 Nov; 65(11): 4476-4482);

Yersinia enterocolitica (gastrointestinal pathogen): particularly LPS (*Infect Immun.* 2002 August; 70(8): 4414);

Yersinia pseudotuberculosis: gastrointestinal pathogen antigens;

Mycobacterium tuberculosis: such as lipoproteins, LPS, BCG antigens, a fusion protein of antigen 85B (Ag85B) and/or ESAT-6 optionally formulated in cationic lipid vesicles (*Infect Immun.* 2004 October; 72(10): 6148), Mycobacterium tuberculosis (Mtb) isocitrate dehydrogenase associated antigens (*Proc Natl Acad Sci U S A.* 2004 Aug 24; 101(34): **12652), and/or MPT51 antigens** (*Infect Immun.* 2004 July; 72(7): 3829);

Legionella pneumophila (Legionnairs' Disease): L. pneumophila antigens -- optionally derived from cell lines with disrupted *asd* genes (*Infect Immun.* 1998 May; 66(5): 1898);

Rickettsia: including outer membrane proteins, including the outer membrane protein A and/or B (OmpB) (*Biochim Biophys Acta.* 2004 Nov 1;1702(2):145), LPS, and surface protein antigen (SPA) (*J Autoimmun.* 1989 Jun;2 Suppl:81);

E. coli: including antigens from enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAggEC), diffusely adhering *E. coli* (DAEC), enteropathogenic *E. coli* (EPEC), and/or enterohemorrhagic *E. coli* (EHEC);

Vibrio cholerae: including proteinase antigens, LPS, particularly lipopolysaccharides of Vibrio cholerae II, O1 Inaba O-specific polysaccharides, V. cholera O139, antigens of IEM108 vaccine (*Infect Immun.* 2003 Oct;71(10):5498-504), and/or Zonula occludens toxin (Zot);

Salmonella typhi (typhoid fever): including capsular polysaccharides preferably conjugates (Vi, i.e. vax-TyVi);

Salmonella typhimurium (gastroenteritis): antigens derived therefrom are contemplated for microbial and cancer therapies, including angiogenesis inhibition and modulation of flk;

Listeria monocytogenes (sytemic infections in immunocompromised or elderly people, infections of fetus): antigens derived from L. monocytogenes are preferably used as carriers/vectors for intracytoplasmic delivery of conjugates/associated compositions of the present invention;

Porphyromonas gingivalis: particularly, *P. gingivalis* outer membrane protein (OMP);

Tetanus: such as tetanus toxoid (TT) antigens, preferably used as a carrier protein in conjunction/conjugated with the compositions of the present invention;

Diphtheria: such as a diphtheria toxoid, preferably CRM_{197}, additionally antigens capable of modulating, inhibiting or associated with ADP ribosylation are contemplated for combination/co-administration/conjugation with the compositions of the present invention, the diphtheria toxoids are preferably used as carrier proteins;

Borrelia burgdorferi (Lyme disease): such as antigens associated with P39 and P13 (an integral membrane protein, *Infect Immun.* 2001 May; 69(5): 3323-3334), VlsE Antigenic Variation Protein (*J Clin Microbiol.* 1999 Dec; 37(12): 3997);

Haemophilus influenzae B: such as a saccharide antigen therefrom;

Klebsiella: such as an OMP, including OMP A, or a polysaccharide optionally conjugated to tetanus toxoid;

Neiserria gonorrhoeae: including, a Por (or porin) protein, such as PorB (*see* Zhu *et al.*, Vaccine (2004) 22:660 – 669), a transferring binding protein, such as TbpA and TbpB (See Price *et al.*, Infection and Immunity (2004) 71(1):277 – 283), a opacity protein (such as Opa), a reduction-modifiable protein (Rmp), and outer membrane vesicle (OMV) preparations (*see* Plante *et al.*, J Infectious Disease (2000) 182:848 – 855), also see *e.g.* WO99/24578, WO99/36544, WO99/57280, WO02/079243);

Chlamydia pneumoniae: particularly *C. pneumoniae* protein antigens;

Chlamydia trachomatis: including antigens derived from serotypes A, B, Ba and C are (agents of trachoma, a cause of blindness), serotypes L_1, L_2 & L_3 (associated with Lymphogranuloma venereum), and serotypes, D-K;

Treponema pallidum (Syphilis): particularly a TmpA antigen; and

Haemophilus ducreyi (causing chancroid): including outer membrane protein (DsrA).

Where not specifically referenced, further bacterial antigens of the invention may be capsular antigens, polysaccharide antigens or protein antigens of any of the above. Further bacterial antigens may also include an outer membrane vesicle (OMV) preparation. Additionally, antigens include live, attenuated, split, and/or purified versions of any of the aforementioned bacteria. The bacterial or microbial derived antigens of the present invention may be gram-negative or gram-positive and aerobic or anaerobic.

Additionally, any of the above bacterial-derived saccharides (polysaccharides, LPS, LOS or oligosaccharides) can be conjugated to another agent or antigen, such as a carrier protein (for example CRM_{197}). Such conjugation may be direct conjugation effected by reductive amination of carbonyl moieties on the saccharide to amino groups on the protein, as provided in US Patent No. 5,360,897 and *Can J Biochem Cell Biol.* 1984 May;62(5):270-5. Alternatively, the saccharides can be conjugated through a linker, such as, with succinamide or other linkages provided in *Bioconjugate Techniques,* 1996 and *CRC, Chemistry of Protein Conjugation and Cross-Linking*, 1993.

Viral Antigens

Influenza: including whole viral particles (attenuated), split, or subunit comprising hemagglutinin (HA) and/or neuraminidase (NA) surface proteins, the influenza antigens may be derived from chicken embryos or propogated on cell culture, and/or the influenza antigens are derived from influenza type A, B, and/or C, among others;

Respiratory syncytial virus (RSV): including the F protein of the A2 strain of RSV (*J Gen Virol.* 2004 Nov; 85(Pt 11):3229) and/or G glycoprotein;

Parainfluenza virus (PIV): including PIV type 1, 2, and 3, preferably containing hemagglutinin, neuraminidase and/or fusion glycoproteins;

Poliovirus: including antigens from a family of picornaviridae, preferably poliovirus antigens such as OPV or, preferably IPV;

Measles: including split measles virus (MV) antigen optionally combined with the Protollin and or antigens present in MMR vaccine;

Mumps: including antigens present in MMR vaccine;

Rubella: including antigens present in MMR vaccine as well as other antigens from Togaviridae, including dengue virus;

Rabies: such as lyophilized inactivated virus (RabAvert™);

Flaviviridae viruses: such as (and antigens derived therefrom) yelow fever virus, Japanese encephalitis virus, dengue virus (types 1, 2, 3, or 4), tick borne encephalitis virus, and West Nile virus;

Caliciviridae; antigens therefrom;

HIV: including HIV-1 or HIV-2 strain antigens, such as gag (p24gag and p55gag), env (gp160 and gp41), pol, tat, nef, rev vpu, miniproteins, (preferably p55 gag and gp140v delete) and antigens from the isolates HIV_{IIIb}, HIV_{SF2}, HIV_{LAV}, HIV_{LAI}, HIV_{MN}, HIV-1_{CM235}, HIV-1_{US4}, HIV-2; simian immunodeficiency virus (SIV) among others;

Rotavirus: including VP4, VP5, VP6, VP7, VP8 proteins (*Protein Expr Purif.* 2004 Dec;38(2):205) and/or NSP4;

Pestivirus: such as antigens from classical porcine fever virus, bovine viral diarrhoea virus, and/or border disease virus;

Parvovirus: such as parvovirus B19;

Coronavirus: including SARS virus antigens, particularly spike protein or proteases therefrom, as well as antigens included in WO 04/92360;

Hepatitis A virus: such as inactivated virus;

Hepatitis B virus: such as the surface and/or core antigens (sAg), as well as the presurface sequences, pre-S1 and pre-S2 (formerly called pre-S), as well as combinations of the above, such as sAg/pre-S1, sAg/pre-S2, sAg/pre-S1/pre-S2, and pre-S1/pre-S2, (see, e.g., AHBV Vaccines - *Human Vaccines and Vaccination*, pp. 159-176; and U.S. Patent Nos. 4,722,840, 5,098,704, 5,324,513;

Beames et al., *J. Virol.* (1995) 69:6833-6838, Birnbaum et al., *J. Virol.* (1990) 64:3319-3330; and Zhou et al., *J. Virol.* (1991) 65:5457-5464);

Hepatitis C virus: such as E1, E2, E1/E2 (see, Houghton et al., *Hepatology* (1991) 14:381), NS345 polyprotein, NS 345-core polyprotein, core, and/or peptides from the nonstructural regions (International Publication Nos. WO 89/04669; WO 90/11089; and WO 90/14436);

Delta hepatitis virus (HDV): antigens derived therefrom, particularly δ-antigen from HDV (see, e.g., U.S. Patent No. 5,378,814);

Hepatitis E virus (HEV); antigens derived therefrom;

Hepatitis G virus (HGV); antigens derived therefrom;

Varcicella zoster virus: antigens derived from varicella zoster virus (VZV) (*J. Gen. Virol.* (1986) 67:1759);

Epstein-Barr virus: antigens derived from EBV (Baer et al., *Nature* (1984) 310:207);

Cytomegalovirus: CMV antigens, including gB and gH (*Cytomegaloviruses* (J.K. McDougall, ed., Springer-Verlag 1990) pp. 125-169);

Herpes simplex virus: including antigens from HSV-1 or HSV-2 strains and glycoproteins gB, gD and gH (McGeoch et al., *J. Gen. Virol.* (1988) 69:1531 and U.S. Patent No. 5,171,568);

Human Herpes Virus: antigens derived from other human herpesviruses such as HHV6 and HHV7; and

HPV: including antigens associated with or derived from human papillomavirus (HPV), for example, one or more of E1 – E7, L1, L2, and fusions thereof, particularly the compositions of the invention may include a virus-like particle (VLP) comprising the L1 major capsid protein, more particular still, the HPV antigens are protective against one or more of HPV serotypes 6, 11, 16 and/or 18.

Further provided are antigens, compostions, methods, and microbes included in *Vaccines*, 4th Edition (Plotkin and Orenstein ed. 2004); *Medical Microbiology* 4th Edition (Murray et al. ed. 2002); *Virology*, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991), which are contemplated in conjunction with the compositions of the present invention.

Additionally, antigens include live, attenuated, split, and/or purified versions of any of the aforementioned viruses.

<u>*Fungal Antigens*</u>

Fungal antigens for use herein, associated with vaccines include those described in: U.S. Pat. Nos. 4,229,434 and 4,368,191 for prophylaxis and treatment of trichopytosis caused by Trichophyton mentagrophytes; U.S. Pat. Nos. 5,277,904 and 5,284,652 for a broad spectrum dermatophyte vaccine for the prophylaxis of dermatophyte infection in animals, such as guinea pigs, cats, rabbits, horses and lambs, these antigens comprises a suspension of killed *T. equinum,* T. mentagrophytes (var. granulare), *M. canis* and/or *M. gypseum* in an effective amount optionally combined with an adjuvant;

U.S. Pat. Nos. 5,453,273 and 6,132,733 for a ringworm vaccine comprising an effective amount of a homogenized, formaldehyde-killed fungi, i.e., *Microsporum canis* culture in a carrier; U.S. Pat. No. 5,948,413 involving extracellular and intracellular proteins for pythiosis. Additional antigens identified within antifungal vaccines include Ringvac bovis LTF-130 and Bioveta.

Further, fungal antigens for use herein may be derived from Dermatophytres, including: *Epidermophyton floccusum, Microsporum audouini, Microsporum canis, Microsporum distortum, Microsporum equinum, Microsporum gypsum, Microsporum nanum, Trichophyton concentricum, Trichophyton equinum, Trichophyton gallinae, Trichophyton gypseum, Trichophyton megnini, Trichophyton mentagrophytes, Trichophyton quinckeanum, Trichophyton rubrum, Trichophyton schoenleini, Trichophyton tonsurans, Trichophyton verrucosum, T. verrucosum* var. album, var. discoides, var. ochraceum, *Trichophyton violaceum,* and/or *Trichophyton faviforme.*

Fungal pathogens for use as antigens or in derivation of antigens in conjunction with the compositions of the present invention comprise *Aspergillus fumigatus, Aspergillus flavus, Aspergillus niger, Aspergillus nidulans, Aspergillus terreus, Aspergillus sydowi, Aspergillus flavatus, Aspergillus glaucus, Blastoschizomyces capitatus, Candida albicans, Candida enolase, Candida tropicalis, Candida glabrata, Candida krusei, Candida parapsilosis, Candida stellatoidea, Candida kusei, Candida parakwsei, Candida lusitaniae, Candida pseudotropicalis, Candida guilliermondi, Cladosporium carrionii, Coccidioides immitis, Blastomyces dermatidis, Cryptococcus neoformans, Geotrichum clavatum, Histoplasma capsulatum, Klebsiella pneumoniae, Paracoccidioides brasiliensis, Pneumocystis carinii, Pythiumn insidiosum, Pityrosporum ovale, Sacharomyces cerevisae, Saccharomyces boulardii, Saccharomyces pombe, Scedosporium apiosperum, Sporothrix schenckii, Trichosporon beigelii, Toxoplasma gondii, Penicillium marneffei,* Malassezia spp., Fonsecaea spp., Wangiella spp., Sporothrix spp., Basidiobolus spp., Conidiobolus spp., Rhizopus spp, Mucor spp, Absidia spp, Mortierella spp, Cunninghamella spp, and Saksenaea spp.

Other fungi from which antigens are derived include Alternaria spp, Curvularia spp, Helminthosporium spp, Fusarium spp, Aspergillus spp, Penicillium spp, Monolinia spp, Rhizoctonia spp, Paecilomyces spp, Pithomyces spp, and Cladosporium spp.

Processes for producing a fungal antigens are well known in the art (see US Patent No. 6,333,164). In a preferred method a solubilized fraction extracted and separated from an insoluble fraction obtainable from fungal cells of which cell wall has been substantially removed or at least partially removed, characterized in that the process comprises the steps of: obtaining living fungal cells; obtaining fungal cells of which cell wall has been substantially removed or at least partially removed; bursting the fungal cells of which cell wall has been substantially removed or at least partially removed; obtaining an insoluble fraction; and extracting and separating a solubilized fraction from the insoluble fraction.

STD Antigens

In particular embodiments, microbes (bacteria, viruses and/or fungi) against which the present compositions and methods can be implement include those that cause sexually transmitted diseases (STDs) and/or those that display on their surface an antigen that can be the target or antigen composition of the invention. In a preferred embodiment of the invention, compositions are combined with antigens derived from a viral or bacterial STD. Antigens derived from bacteria or viruses can be administered in conjunction with the compositions of the present invention to provide protection against at least one of the following STDs, among others: chlamydia, genital herpes, hepatitis (particularly HCV), genital warts, gonorrhoea, syphilis and/or chancroid (See, WO00/15255).

In another embodiment the compositions of the present invention are co-administered with an antigen for the prevention or treatment of an STD.

Antigens derived from the following viruses associated with STDs, which are described in greater detail above, are preferred for co-administration with the compositions of the present invention: hepatitis (particularly HCV), HPV, HIV, or HSV.

Additionally, antigens derived from the following bacteria associated with STDs, which are described in greater detail above, are preferred for co-administration with the compositions of the present invention: *Neiserria gonorrhoeae, Chlamydia pneumoniae, Chlamydia trachomatis, Treponema pallidum*, or *Haemophilus ducreyi.*

Respiratory Antigens

The antigen may be a respiratory antigen and could further be used in an immunogenic composition for methods of preventing and/or treating infection by a respiratory pathogen, including a virus, bacteria, or fungi such as respiratory syncytial virus (RSV), PIV, SARS virus, influenza, *Bacillus anthracis,* particularly by reducing or preventing infection and/or one or more symptoms of respiratory virus infection. A composition comprising an antigen described herein, such as one derived from a respiratory virus, bacteria or fungus is administered in conjunction with the compositions of the present invention to an individual which is at risk of being exposed to that particular respiratory microbe, has been exposed to a respiratory microbe or is infected with a respiratory virus, bacteria or fungus. The composition(s) of the present invention is/are preferably co-administered at the same time or in the same formulation with an antigen of the respiratory pathogen. Administration of the composition results in reduced incidence and/or severity of one or more symptoms of respiratory infection.

Pediatric/Geriatric Antigens

In one embodiment the compositions of the present invention are used in conjunction with an antigen for treatment of a pediatric population, as in a pediatric antigen. In a more particular embodiment the pediatric population is less than about 3 years old, or less than about 2 years, or less than about 1 years old. In another embodiment the pediatric antigen (in conjunction with the composition of the present invention) is administered multiple times over at least 1, 2, or 3 years.

In another embodiment the compositions of the present invention are used in conjunction with an antigen for treatment of a geriatric population, as in a geriatric antigen.

Other Antigens

Other antigens for use in conjunction with the compositions of the present include hospital acquired (nosocomial) associated antigens.

In another embodiment, parasitic antigens are contemplated in conjunction with the compositions of the present invention. Examples of parasitic antigens include those derived from organisms causing malaria and/or Lyme disease.

In another embodiment, the antigens in conjunction with the compositions of the present invention are associated with or effective against a mosquito born illness. In another embodiment, the antigens in conjunction with the compositions of the present invention are associated with or effective against encephalitis. In another embodiment the antigens in conjunction with the compositions of the present invention are associated with or effective against an infection of the nervous system.

In another embodiment, the antigens in conjunction with the compositions of the present invention are antigens transmissible through blood or body fluids.

Antigen Formulations

In other aspects of the invention, methods of producing microparticles having adsorbed antigens are provided. The methods comprise: (a) providing an emulsion by dispersing a mixture comprising (i) water, (ii) a detergent, (iii) an organic solvent, and (iv) a biodegradable polymer selected from the group consisting of a poly(α-hydroxy acid), a polyhydroxy butyric acid, a polycaprolactone, a polyorthoester, a polyanhydride, and a polycyanoacrylate. The polymer is typically present in the mixture at a concentration of about 1% to about 30% relative to the organic solvent, while the detergent is typically present in the mixture at a weight-to-weight detergent-to-polymer ratio of from about 0.00001:1 to about 0.1:1 (more typically about 0.0001:1 to about 0.1:1, about 0.001:1 to about 0.1:1, or about 0.005:1 to about 0.1:1); (b) removing the organic solvent from the emulsion; and (c) adsorbing an antigen on the surface of the microparticles. In certain embodiments, the biodegradable polymer is present at a concentration of about 3% to about 10% relative to the organic solvent.

Microparticles for use herein will be formed from materials that are sterilizable, non-toxic and biodegradable. Such materials include, without limitation, poly(α-hydroxy acid), polyhydroxybutyric acid, polycaprolactone, polyorthoester, polyanhydride, PACA, and polycyanoacrylate. Preferably, microparticles for use with the present invention are derived from a poly(α-hydroxy acid), in particular, from a poly(lactide) ("PLA") or a copolymer of D,L-lactide and glycolide or glycolic acid, such as a poly(D,L-lactide-co-glycolide) ("PLG" or "PLGA"), or a copolymer of D,L-lactide and caprolactone. The microparticles may be derived from any of various polymeric starting materials which have a variety of molecular weights and, in the case of the copolymers such as PLG, a variety of lactide:glycolide ratios, the selection of which will be largely a

matter of choice, depending in part on the coadministered macromolecule. These parameters are discussed more fully below.

Further antigens may also include an outer membrane vesicle (OMV) preparation.

Additional formulation methods and antigens (especially tumor antigens) are provided in U.S. Patent Serial No. 09/581,772.

Antigen References

The following references include antigens useful in conjunction with the compositions of the present invention:

1 International patent application WO99/24578
2 International patent application WO99/36544.
3 International patent application WO99/57280.
4 International patent application WO00/22430.
5 Tettelin et al. (2000) Science 287:1809-1815.
6 International patent application WO96/29412.
7 Pizza et al. (2000) Science 287:1816-1820.
8 PCT WO 01/52885.
9 Bjune et al. (1991) Lancet 338(8775).
10 Fuskasawa et al. (1999) Vaccine 17:2951-2958.
11 Rosenqist et al. (1998) Dev. Biol. Strand 92:323-333.
12 Constantino et al. (1992) Vaccine 10:691-698.
13 Constantino et al. (1999) Vaccine 17:1251-1263.
14 Watson (2000) Pediatr Infect Dis J 19:331-332.
15 Rubin (20000) Pediatr Clin North Am 47:269-285,v.
16 Jedrzejas (2001) Microbiol Mol Biol Rev 65:187-207.
17 International patent application filed on 3rd July 2001 claiming priority from GB-0016363.4;WO 02/02606; PCT IB/01/00166.
18 Kalman et al. (1999) Nature Genetics 21:385-389.
19 Read et al. (2000) Nucleic Acids Res 28:1397-406.
20 Shirai et al. (2000) J. Infect. Dis 181(Suppl 3):S524-S527.
21 International patent application WO99/27105.
22 International patent application WO00/27994.
23 International patent application WO00/37494.
24 International patent application WO99/28475.
25 Bell (2000) Pediatr Infect Dis J 19:1187-1188.
26 Iwarson (1995) APMIS 103:321-326.
27 Gerlich et al. (1990) Vaccine 8 Suppl:S63-68 & 79-80.
28 Hsu et al. (1999) Clin Liver Dis 3:901-915.
29 Gastofsson et al. (1996) N. Engl. J. Med. 334-:349-355.
30 Rappuoli et al. (1991) TIBTECH 9:232-238.
31 Vaccines (1988) eds. Plotkin & Mortimer. ISBN 0-7216-1946-0.
32 Del Guidice et al. (1998) Molecular Aspects of Medicine 19:1-70.
33 International patent application WO93/018150.
34 International patent application WO99/53310.
35 International patent application WO98/04702.
36 Ross et al. (2001) Vaccine 19:135-142.
37 Sutter et al. (2000) Pediatr Clin North Am 47:287-308.
38 Zimmerman & Spann (1999) Am Fan Physician 59:113-118, 125-126.
39 Dreensen (1997) Vaccine 15 Suppl"S2-6.
40 MMWR Morb Mortal Wkly rep 1998 Jan 16:47(1):12, 9.
41 McMichael (2000) Vaccine19 Suppl 1:S101-107.

42 Schuchat (1999) Lancet 353(9146):51-6.
43 GB patent applications 0026333.5, 0028727.6 & 0105640.7.
44 Dale (1999) Infect Disclin North Am 13:227-43, viii.
45 Ferretti et al. (2001) PNAS USA 98: 4658-4663.
46 Kuroda et al. (2001) Lancet 357(9264):1225-1240; see also pages 1218-1219.
47 Ramsay et al. (2001) Lancet 357(9251):195-196.
48 Lindberg (1999) Vaccine 17 Suppl 2:S28-36.
49 Buttery & Moxon (2000) J R Coil Physicians Long 34:163-168.
50 Ahmad & Chapnick (1999) Infect Dis Clin North Am 13:113-133, vii.
51 Goldblatt (1998) J. Med. Microbiol. 47:663-567.
52 European patent 0 477 508.
53 U.S. Patent No. 5,306,492.
54 International patent application WO98/42721.
55 Conjugate Vaccines (eds. Cruse et al.) ISBN 3805549326, particularly vol. 10:48-114.
56 Hermanson (1996) Bioconjugate Techniques ISBN: 012323368 & 012342335X.
57 European patent application 0372501.
58 European patent application 0378881.
59 European patent application 0427347.
60 International patent application WO93/17712.
61 International patent application WO98/58668.
62 European patent application 0471177.
63 International patent application WO00/56360.
64 International patent application WO00/67161.

The contents of all of the above cited patents, patent applications and journal articles are incorporated by reference as if set forth fully herein.

There may be an upper limit to the number of Gram positive bacterial proteins which will be in the compositions of the invention. Preferably, the number of Gram positive bacterial proteins in a composition of the invention is less than 20, less than 19, less than 18, less than 17, less than 16, less than 15, less than 14, less than 13, less than 12, less than 11, less than 10, less than 9, less than 8, less than 7, less than 6, less than 5, less than 4, or less than 3. Still more preferably, the number of Gram positive bacterial proteins in a composition of the invention is less than 6, less than 5, or less than 4. Still more preferably, the number of Gram positive bacterial proteins in a composition of the invention is 3.

The Gram positive bacterial proteins and polynucleotides used in the invention are preferably isolated, *i.e.*, separate and discrete, from the whole organism with which the molecule is found in nature or, when the polynucleotide or polypeptide is not found in nature, is sufficiently free of other biological macromolecules so that the polynucleotide or polypeptide can be used for its intended purpose.

Fusion Proteins: GBS AI sequences

The GBS AI proteins used in the invention may be present in the composition as individual separate polypeptides, but it is preferred that at least two (*i.e.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18) of the antigens are expressed as a single polypeptide chain (a "hybrid" or "fusion" polypeptide). Such fusion polypeptides offer two principal advantages: first, a polypeptide that may be unstable or poorly expressed on its own can be assisted by adding a suitable fusion partner that

overcomes the problem; second, commercial manufacture is simplified as only one expression and purification need be employed in order to produce two polypeptides which are both antigenically useful.

The fusion polypeptide may comprise one or more AI polypeptide sequences. Preferably, the fusion comprises an AI surface protein sequence. Preferably, the fusion polypeptide includes one or more of GBS 80, GBS 104, and GBS 67. Most preferably, the fusion peptide includes a polypeptide sequence from GBS 80. Accordingly, the invention includes a fusion peptide comprising a first amino acid sequence and a second amino acid sequence, wherein said first and second amino acid sequences are selected from a GBS AI surface protein or a fragment thereof. Preferably, the first and second amino acid sequences in the fusion polypeptide comprise different epitopes.

Hybrids (or fusions) consisting of amino acid sequences from two, three, four, five, six, seven, eight, nine, or ten GBS antigens are preferred. In particular, hybrids consisting of amino acid sequences from two, three, four, or five GBS antigens are preferred.

Different hybrid polypeptides may be mixed together in a single formulation. Within such combinations, a GBS antigen may be present in more than one hybrid polypeptide and/or as a non-hybrid polypeptide. It is preferred, however, that an antigen is present either as a hybrid or as a non-hybrid, but not as both.

Hybrid polypeptides can be represented by the formula $\mathbf{NH_2\text{-}A\text{-}\{\text{-}X\text{-}L\text{-}\}_{\mathit{n}}\text{-}B\text{-}COOH}$, wherein: X is an amino acid sequence of a GBS AI protein or a fragment thereof; L is an optional linker amino acid sequence; A is an optional N-terminal amino acid sequence; B is an optional C-terminal amino acid sequence; and *n* is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15.

If a -X- moiety has a leader peptide sequence in its wild-type form, this may be included or omitted in the hybrid protein. In some embodiments, the leader peptides will be deleted except for that of the -X- moiety located at the N-terminus of the hybrid protein *i.e.* the leader peptide of X_1 will be retained, but the leader peptides of $X_2 \ldots X_n$ will be omitted. This is equivalent to deleting all leader peptides and using the leader peptide of X_1 as moiety -A-.

For each *n* instances of {-X-L-}, linker amino acid sequence -L- may be present or absent. For instance, when *n*=2 the hybrid may be NH_2-X_1-L_1-X_2-L_2-COOH, NH_2-X_1-X_2-COOH, NH_2-X_1-L_1-X_2-COOH, NH_2-X_1-X_2-L_2-COOH, *etc.* Linker amino acid sequence(s) -L- will typically be short (*e.g.* 20 or fewer amino acids *i.e.* 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples comprise short peptide sequences which facilitate cloning, poly-glycine linkers (*i.e.* comprising Gly_n where n = 2, 3, 4, 5, 6, 7, 8, 9, 10 or more), and histidine tags (*i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable linker amino acid sequences will be apparent to those skilled in the art. A useful linker is GSGGGG, with the Gly-Ser dipeptide being formed from a *Bam*HI restriction site, thus aiding cloning and manipulation, and the $(Gly)_4$ tetrapeptide being a typical poly-glycine linker.

-A- is an optional N-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19,

18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include leader sequences to direct protein trafficking, or short peptide sequences which facilitate cloning or purification (*e.g.* histidine tags *i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable N-terminal amino acid sequences will be apparent to those skilled in the art. If X_1 lacks its own N-terminus methionine, -A- is preferably an oligopeptide (*e.g.* with 1, 2, 3, 4, 5, 6, 7 or 8 amino acids) which provides a N-terminus methionine.

-B- is an optional C-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include sequences to direct protein trafficking, short peptide sequences which facilitate cloning or purification (*e.g.* comprising histidine tags *i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more), or sequences which enhance protein stability. Other suitable C-terminal amino acid sequences will be apparent to those skilled in the art.

Most preferably, n is 2 or 3.

Fusion Proteins: Gram positive bacteria AI sequences

The Gram positive bacteria AI proteins used in the invention may be present in the composition as individual separate polypeptides, but it is preferred that at least two (*i.e.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18) of the antigens are expressed as a single polypeptide chain (a "hybrid" or "fusion" polypeptide). Such fusion polypeptides offer two principal advantages: first, a polypeptide that may be unstable or poorly expressed on its own can be assisted by adding a suitable fusion partner that overcomes the problem; second, commercial manufacture is simplified as only one expression and purification need be employed in order to produce two polypeptides which are both antigenically useful.

The fusion polypeptide may comprise one or more AI polypeptide sequences. Preferably, the fusion comprises an AI surface protein sequence. Accordingly, the invention includes a fusion peptide comprising a first amino acid sequence and a second amino acid sequence, wherein said first and second amino acid sequences are selected from a Gram positive bacteria AI protein or a fragment thereof. Preferably, the first and second amino acid sequences in the fusion polypeptide comprise different epitopes.

Hybrids (or fusions) consisting of amino acid sequences from two, three, four, five, six, seven, eight, nine, or ten Gram positive bacteria antigens are preferred. In particular, hybrids consisting of amino acid sequences from two, three, four, or five Gram positive bacteria antigens are preferred.

Different hybrid polypeptides may be mixed together in a single formulation. Within such combinations, a Gram positive bacteria AI sequence may be present in more than one hybrid polypeptide and/or as a non-hybrid polypeptide. It is preferred, however, that an antigen is present either as a hybrid or as a non-hybrid, but not as both.

Hybrid polypeptides can be represented by the formula **NH_2-A-{-X-L-}$_n$-B-COOH**, wherein: X is an amino acid sequence of a Gram positive bacteria AI sequence or a fragment thereof; L is an

optional linker amino acid sequence; A is an optional N-terminal amino acid sequence; B is an optional C-terminal amino acid sequence; and *n* is 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15.

If a -X- moiety has a leader peptide sequence in its wild-type form, this may be included or omitted in the hybrid protein. In some embodiments, the leader peptides will be deleted except for that of the -X- moiety located at the N-terminus of the hybrid protein *i.e.* the leader peptide of X_1 will be retained, but the leader peptides of $X_2 \ldots X_n$ will be omitted. This is equivalent to deleting all leader peptides and using the leader peptide of X_1 as moiety -A-.

For each *n* instances of {-X-L-}, linker amino acid sequence -L- may be present or absent. For instance, when n=2 the hybrid may be NH_2-X_1-L_1-X_2-L_2-COOH, NH_2-X_1-X_2-COOH, NH_2-X_1-L_1-X_2-COOH, NH_2-X_1-X_2-L_2-COOH, *etc.* Linker amino acid sequence(s) -L- will typically be short (*e.g.* 20 or fewer amino acids *i.e.* 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples comprise short peptide sequences which facilitate cloning, poly-glycine linkers (*i.e.* comprising Gly_n where n = 2, 3, 4, 5, 6, 7, 8, 9, 10 or more), and histidine tags (*i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable linker amino acid sequences will be apparent to those skilled in the art. A useful linker is GSGGGG, with the Gly-Ser dipeptide being formed from a *Bam*HI restriction site, thus aiding cloning and manipulation, and the $(Gly)_4$ tetrapeptide being a typical poly-glycine linker.

-A- is an optional N-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include leader sequences to direct protein trafficking, or short peptide sequences which facilitate cloning or purification (*e.g.* histidine tags *i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more). Other suitable N-terminal amino acid sequences will be apparent to those skilled in the art. If X_1 lacks its own N-terminus methionine, -A- is preferably an oligopeptide (*e.g.* with 1, 2, 3, 4, 5, 6, 7 or 8 amino acids) which provides a N-terminus methionine.

-B- is an optional C-terminal amino acid sequence. This will typically be short (*e.g.* 40 or fewer amino acids *i.e.* 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Examples include sequences to direct protein trafficking, short peptide sequences which facilitate cloning or purification (*e.g.* comprising histidine tags *i.e.* His_n where n = 3, 4, 5, 6, 7, 8, 9, 10 or more), or sequences which enhance protein stability. Other suitable C-terminal amino acid sequences will be apparent to those skilled in the art.

Most preferably, *n* is 2 or 3.

Antibodies: GBS AI sequences

The GBS AI proteins of the invention may also be used to prepare antibodies specific to the GBS AI proteins. The antibodies are preferably specific to the an oligomeric or hyper-oligomeric form of an AI protein. The invention also includes combinations of antibodies specific to GBS AI proteins selected to provide protection against an increased range of GBS serotypes and strain isolates. For example, a combination may comprise a first and second antibody, wherein said first

antibody is specific to a first GBS AI protein and said second antibody is specific to a second GBS AI protein. Preferably, the nucleic acid sequence encoding said first GBS AI protein is not present in a GBS genome comprising a polynucleotide sequence encoding for said second GBS AI protein. Preferably, the nucleic acid sequence encoding said first and second GBS AI proteins are present in the genomes of multiple GBS serotypes and strain isolates.

The GBS specific antibodies of the invention include one or more biological moieties that, through chemical or physical means, can bind to or associate with an epitope of a GBS polypeptide. The antibodies of the invention include antibodies which specifically bind to a GBS AI protein. The invention includes antibodies obtained from both polyclonal and monoclonal preparations, as well as the following: hybrid (chimeric) antibody molecules (see, for example, Winter *et al.* (1991) *Nature* 349: 293-299; and US Patent No. 4,816,567; $F(ab')_2$ and F(ab) fragments; F_v molecules (non-covalent heterodimers, see, for example, Inbar *et al.* (1972) *Proc Natl Acad Sci USA* 69:2659-2662; and Ehrlich *et al.* (1980) *Biochem* 19:4091-4096); single-chain Fv molecules (sFv) (see, for example, Huston *et al.* (1988) *Proc Natl Acad Sci USA* 85:5897-5883); dimeric and trimeric antibody fragment constructs; minibodies (see, *e.g.*, Pack *et al.* (1992) *Biochem* 31:1579-1584; Cumber *et al.* (1992) *J Immunology* 149B: 120-126); humanized antibody molecules (see, for example, Riechmann *et al.* (1988) *Nature* 332:323-327; Verhoeyan *et al.* (1988) *Science* 239:1534-1536; and U.K. Patent Publication No. GB 2,276,169, published 21 September 1994); and, any functional fragments obtained from such molecules, wherein such fragments retain immunological binding properties of the parent antibody molecule. The invention further includes antibodies obtained through non-conventional processes, such as phage display.

Preferably, the GBS specific antibodies of the invention are monoclonal antibodies. Monoclonal antibodies of the invention include an antibody composition having a homogeneous antibody population. Monoclonal antibodies of the invention may be obtained from murine hybridomas, as well as human monoclonal antibodies obtained using human rather than murine hybridomas. See, *e.g.*, Cote, *et al. Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 1985, p 77.

The antibodies of the invention may be used in diagnostic applications, for example, to detect the presence or absence of GBS in a biological sample. The antibodies of the invention may also be used in the prophylactic or therapeutic treatment of GBS infection.

Antibodies: Gram positive bacteria AI sequences

The Gram positive bacteria AI proteins of the invention may also be used to prepare antibodies specific to the Gram positive bacteria AI proteins. The antibodies are preferably specific to the an oligomeric or hyper-oligomeric form of an AI protein. The invention also includes combinations of antibodies specific to Gram positive bacteria AI proteins selected to provide protection against an increased range of Gram positive bacteria genus, species, serotypes and strain isolates.

For example, a combination may comprise a first and second antibody, wherein said first antibody is specific to a first Gram positive bacteria AI protein and said second antibody is specific to a second Gram positive bacteria AI protein. Preferably, the nucleic acid sequence encoding said first Gram positive bacteria AI protein is not present in a Gram positive bacterial genome comprising a polynucleotide sequence encoding for said second Gram positive bacteria AI protein. Preferably, the nucleic acid sequence encoding said first and second Gram positive bacteria AI proteins are present in the genomes of multiple Gram positive bacteria genus, species, serotypes or strain isolates.

As an example of an instance where the combination of antibodies provides protection against an increased range of bacteria serotypes, the first antibody may be specific to a first GAS AI protein and the second antibody may be specific to a second GAS AI protein. The first GAS AI protein may comprise a GAS AI-1 surface protein, while the second GAS AI protein may comprise a GAS AI-2 or AI-3 surface protein.

As an example of an instance where the combination of antibodies provides protection against an increased range of bacterial species, the first antibody may be specific to a GBS AI protein and the second antibody may be specific to a GAS AI protein. Alternatively, the first antibody may be specific to a GAS AI protein and the second antibody may be specific to a *S. pneumoniae* AI protein.

The Gram positive specific antibodies of the invention include one or more biological moieties that, through chemical or physical means, can bind to or associate with an epitope of a Gram positive bacteria AI polypeptide. The antibodies of the invention include antibodies which specifically bind to a Gram positive bacteria AI protein. The invention includes antibodies obtained from both polyclonal and monoclonal preparations, as well as the following: hybrid (chimeric) antibody molecules (see, for example, Winter *et al.* (1991) *Nature* 349: 293-299; and US Patent No. 4,816,567; $F(ab')_2$ and F(ab) fragments; F_v molecules (non-covalent heterodimers, see, for example, Inbar *et al.* (1972) *Proc Natl Acad Sci USA* 69:2659-2662; and Ehrlich *et al.* (1980) *Biochem* 19:4091-4096); single-chain Fv molecules (sFv) (see, for example, Huston *et al.* (1988) *Proc Natl Acad Sci USA* 85:5897-5883); dimeric and trimeric antibody fragment constructs; minibodies (see, *e.g.*, Pack *et al.* (1992) *Biochem* 31:1579-1584; Cumber *et al.* (1992) *J Immunology* 149B: 120-126); humanized antibody molecules (see, for example, Riechmann *et al.* (1988) *Nature* 332:323-327; Verhoeyan *et al.* (1988) *Science* 239:1534-1536; and U.K. Patent Publication No. GB 2,276,169, published 21 September 1994); and, any functional fragments obtained from such molecules, wherein such fragments retain immunological binding properties of the parent antibody molecule. The invention further includes antibodies obtained through non-conventional processes, such as phage display.

Preferably, the Gram positive specific antibodies of the invention are monoclonal antibodies. Monoclonal antibodies of the invention include an antibody composition having a homogeneous antibody population. Monoclonal antibodies of the invention may be obtained from murine hybridomas, as well as human monoclonal antibodies obtained using human rather than murine

hybridomas. See, e.g., Cote, *et al. Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 1985, p 77.

The antibodies of the invention may be used in diagnostic applications, for example, to detect the presence or absence of Gram positive bacteria in a biological sample. The antibodies of the invention may also be used in the prophylactic or therapeutic treatment of Gram positive bacteria infection.

Nucleic Acids

The invention provides nucleic acids encoding the Gram positive bacteria sequences and/or the hybrid fusion polypeptides of the invention. The invention also provides nucleic acid encoding the GBS antigens and/or the hybrid fusion polypeptides of the invention. Furthermore, the invention provides nucleic acid which can hybridise to these nucleic acids, preferably under "high stringency" conditions (*e.g.* 65°C in a 0.1xSSC, 0.5% SDS solution).

Polypeptides of the invention can be prepared by various means (*e.g.* recombinant expression, purification from cell culture, chemical synthesis, *etc.*) and in various forms (*e.g.* native, fusions, non-glycosylated, lipidated, *etc.*). They are preferably prepared in substantially pure form (*i.e.* substantially free from other GAS or host cell proteins).

Nucleic acid according to the invention can be prepared in many ways (*e.g.* by chemical synthesis, from genomic or cDNA libraries, from the organism itself, *etc.*) and can take various forms (*e.g.* single stranded, double stranded, vectors, probes, *etc.*). They are preferably prepared in substantially pure form (*i.e.* substantially free from other GBS or host cell nucleic acids).

The term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones (*e.g.* phosphorothioates, *etc.*), and also peptide nucleic acids (PNA), *etc.* The invention includes nucleic acid comprising sequences complementary to those described above (*e.g.* for antisense or probing purposes).

The invention also provides a process for producing a polypeptide of the invention, comprising the step of culturing a host cell transformed with nucleic acid of the invention under conditions which induce polypeptide expression.

The invention provides a process for producing a polypeptide of the invention, comprising the step of synthesising at least part of the polypeptide by chemical means.

The invention provides a process for producing nucleic acid of the invention, comprising the step of amplifying nucleic acid using a primer-based amplification method (*e.g.* PCR).

The invention provides a process for producing nucleic acid of the invention, comprising the step of synthesising at least part of the nucleic acid by chemical means.

Purification and Recombinant Expression

The Gram positive bacteria AI proteins of the invention may be isolated from the native Gram positive bacteria, or they may be recombinantly produced, for instance in a heterologous host. For example, the GAS, GBS, and *S. pneumoniae* antigens of the invention may be isolated from

Streptococcus agalactiae, S. pyogenes, S. pneumoniae, or they may be recombinantly produced, for instance, in a heterologous host. Preferably, the GBS antigens are prepared using a heterologous host.

The heterologous host may be prokaryotic (*e.g.* a bacterium) or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis, Vibrio cholerae, Salmonella typhi, Salmonella typhimurium, Neisseria lactamica, Neisseria cinerea, Mycobacteria* (*e.g. M.tuberculosis*), *S. gordonii, L. lactis*, yeasts, *etc.*

Recombinant production of polypeptides is facilitated by adding a tag protein to the Gram positive bacteria AI sequence to be expressed as a fusion protein comprising the tag protein and the Gram positive bacteria antigen. For example, recombinant production of polypeptides is facilitated by adding a tag protein to the GBS antigen to be expressed as a fusion protein comprising the tag protein and the GBS antigen. Such tag proteins can facilitate purification, detection and stability of the expressed protein. Tag proteins suitable for use in the invention include a polyarginine tag (Arg-tag), polyhistidine tag (His-tag), FLAG-tag, Strep-tag, c-myc-tag, S-tag, calmodulin-binding peptide, cellulose-binding domain, SBP-tag,, chitin-binding domain, glutathione S-transferase-tag (GST), maltose-binding protein, transcription termination anti-terminiantion factor (NusA), *E. coli* thioredoxin (TrxA) and protein disulfide isomerase I (DsbA). Preferred tag proteins include His-tag and GST. A full discussion on the use of tag proteins can be found at Terpe et al., "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems", Appl Microbiol Biotechnol (2003) 60:523 – 533.

After purification, the tag proteins may optionally be removed from the expressed fusion protein, *i.e.*, by specifically tailored enzymatic treatments known in the art. Commonly used proteases include enterokinase, tobacco etch virus (TEV), thrombin, and factor X_a.

GBS polysaccharides

The compositions of the invention may be further improved by including GBS polysaccharides. Preferably, the GBS antigen and the saccharide each contribute to the immunological response in a recipient. The combination is particularly advantageous where the saccharide and polypeptide provide protection from different GBS serotypes.

The combined antigens may be present as a simple combination where separate saccharide and polypeptide antigens are administered together, or they may be present as a conjugated combination, where the saccharide and polypeptide antigens are covalently linked to each other.

Thus the invention provides an immunogenic composition comprising (i) one or more GBS AI proteins and (ii) one or more GBS saccharide antigens. The polypeptide and the polysaccharide may advantageously be covalently linked to each other to form a conjugate.

Between them, the combined polypeptide and saccharide antigens preferably cover (or provide protection from) two or more GBS serotypes (*e.g.* 2, 3, 4, 5, 6, 7, 8 or more serotypes). The serotypes of the polypeptide and saccharide antigens may or may not overlap. For example, the polypeptide might protect against serogroup II or V, while the saccharide protects against either serogroups Ia, Ib, or III. Preferred combinations protect against the following groups of serotypes:

(1) serotypes Ia and Ib, (2) serotypes Ia and II, (3) serotypes Ia and III, (4) serotypes Ia and IV, (5) serotypes Ia and V, (6) serotypes Ia and VI, (7) serotypes Ia and VII, (8) serotypes Ia and VIII, (9) serotypes Ib and II, (10) serotypes Ib and III, (11) serotypes Ib and IV, (12) serotypes Ib and V, (13) serotypes Ib and VI, (14) serotypes Ib and VII, (15) serotypes Ib and VIII, 16) serotypes II and III, (17) serotypes II and IV, (18) serotypes II and V, (19) serotypes II and VI, (20) serotypes II and VII, (21) serotypes II and VII, (22) serotypes III and IV, (23) serotypes III and V, (24) serotypes III and VI, (25) serotypes III and VII, (26) serotypes III and VIII, (27) serotypes IV and V, (28) serotypes IV and VI, (29) serotypes IV and VII, (30) serotypes IV and VIII, (31) serotypes V and VI, (32) serotypes V and VII, (33) serotypes V and VIII, (34) serotypes VI and VII, (35) serotypes VI and VIII, and (36) serotypes VII and VIII.

Still more preferably, the combinations protect against the following groups of serotypes: (1) serotypes Ia and II, (2) serotypes Ia and V, (3) serotypes Ib and II, (4) serotypes Ib and V, (5) serotypes III and II, and (6) serotypes III and V. Most preferably, the combinations protect against serotypes III and V.

Protection against serotypes II and V is preferably provided by polypeptide antigens. Protection against serotypes Ia, Ib and/or III may be polypeptide or saccharide antigens.

Immunogenic compositions and medicaments

Compositions of the invention are preferably immunogenic compositions, and are more preferably vaccine compositions. The pH of the composition is preferably between 6 and 8, preferably about 7. The pH may be maintained by the use of a buffer. The composition may be sterile and/or pyrogen-free. The composition may be isotonic with respect to humans.

Vaccines according to the invention may either be prophylactic (*i.e.* to prevent infection) or therapeutic (*i.e.* to treat infection), but will typically be prophylactic. Accordingly, the invention includes a method for the therapeutic or prophylactic treatment of a Gram positive bacteria infection in an animal susceptible to such gram positive bacterial infection comprising administering to said animal a therapeutic or prophylactic amount of the immunogenic composition of the invention. For example, the invention includes a method for the therapeutic or prophylactic treatment of a *Streptococcus agalactiae*, *S. pyogenes*, or *S. pneumoniae* infection in an animal susceptible to streptococcal infection comprising administering to said animal a therapeutic or prophylactic amount of the immunogenic compositions of the invention.

The invention also provides a composition of the invention for use of the compositions described herein as a medicament. The medicament is preferably able to raise an immune response in a mammal (*i.e.* it is an immunogenic composition) and is more preferably a vaccine.

The invention also provides the use of the compositions of the invention in the manufacture of a medicament for raising an immune response in a mammal. The medicament is preferably a vaccine.

The invention also provides kits comprising one or more containers of compositions of the invention. Compositions can be in liquid form or can be lyophilized, as can individual antigens. Suitable containers for the compositions include, for example, bottles, vials, syringes, and test tubes.

Containers can be formed from a variety of materials, including glass or plastic. A container may have a sterile access port (for example, the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The composition may comprise a first component comprising one or more Gram positive bacteria AI proteins. Preferably, the AI proteins are surface AI proteins. Preferably, the AI surface proteins are in an oligomeric or hyperoligomeric form. For example, the first component comprises a combination of GBS antigens or GAS antigens, or *S. pneumoniae* antigens. Preferably said combination includes GBS 80. Preferably GBS 80 is present in an oligomeric or hyperoligomeric form.

The kit can further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, or dextrose solution. It can also contain other materials useful to the end-user, including other buffers, diluents, filters, needles, and syringes. The kit can also comprise a second or third container with another active agent, for example an antibiotic.

The kit can also comprise a package insert containing written instructions for methods of inducing immunity against S agalactiae andor *S. pyogenes* and/or *S pneumoniae* or for treating S agalactiae andor *S. pyogenes* and/or S pneumoniae infections. The package insert can be an unapproved draft package insert or can be a package insert approved by the Food and Drug Administration (FDA) or other regulatory body.

The invention also provides a delivery device pre-filled with the immunogenic compositions of the invention.

The invention also provides a method for raising an immune response in a mammal comprising the step of administering an effective amount of a composition of the invention. The immune response is preferably protective and preferably involves antibodies and/or cell-mediated immunity. This immune response will preferably induce long lasting (*e.g.*, neutralising) antibodies and a cell mediated immunity that can quickly respond upon exposure to one or more GBS and/or GAS and/or *S. pneumoniae* antigens. The method may raise a booster response.

The invention provides a method of neutralizing GBS, GAS, or *S. pneumoniae* infection in a mammal comprising the step of administering to the mammal an effective amount of the immunogenic compositions of the invention, a vaccine of the invention, or antibodies which recognize an immunogenic composition of the invention.

The mammal is preferably a human. Where the vaccine is for prophylactic use, the human is preferably a female (either of child bearing age or a teenager). Alternatively, the human may be elderly (e.g., over the age of 50, 55, 60, 65, 70 or 75) and may have an underlying disease such as diabetes or cancer. Where the vaccine is for therapeutic use, the human is preferably a pregnant female or an elderly adult.

These uses and methods are preferably for the prevention and/or treatment of a disease caused by *Streptococcus agalactiae*, or *S. pyogenes*, or *S. pneumoniae*. The compositions may also be

effective against other streptococcal bacteria. The compositions may also be effective against other Gram positive bacteria.

One way of checking efficacy of therapeutic treatment involves monitoring Gram positive bacterial infection after administration of the composition of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses against the Gram positive bacterial antigens in the compositions of the invention after administration of the composition.

One way of checking efficacy of therapeutic treatment involves monitoring GBS infection after administration of the composition of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses against the GBS antigens in the compositions of the invention after administration of the composition.

A way of assessing the immunogenicity of the component proteins of the immunogenic compositions of the present invention is to express the proteins recombinantly and to screen patient sera or mucosal secretions by immunoblot. A positive reaction between the protein and the patient serum indicates that the patient has previously mounted an immune response to the protein in question- that is, the protein is an immunogen. This method may also be used to identify immunodominant proteins and/or epitopes.

Another way of checking efficacy of therapeutic treatment involves monitoring GBS or GAS or *S pneumoniae* infection after administration of the compositions of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses both systemically (such as monitoring the level of IgG1 and IgG2a production) and mucosally (such as monitoring the level of IgA production) against the GBS and/or GAS and/or *S pneumoniae* antigens in the compositions of the invention after administration of the composition. Typically, GBS and/or GAS and/or S pneumoniae serum specific antibody responses are determined post-immunization but pre-challenge whereas mucosal GBS and/or GAS and/or *S pneumoniae* specific antibody body responses are determined post-immunization and post-challenge.

The vaccine compositions of the present invention can be evaluated in *in vitro* and *in vivo* animal models prior to host, *e.g.*, human, administration.

The efficacy of immunogenic compositions of the invention can also be determined *in vivo* by challenging animal models of GBS and/or GAS and/or S pneumoniae infection, *e.g.*, guinea pigs or mice, with the immunogenic compositions. The immunogenic compositions may or may not be derived from the same serotypes as the challenge serotypes. Preferably the immunnogenic compositions are derivable from the same serotypes as the challenge serotypes. More preferably, the immunogenic composition and/or the challenge serotypes are derivable from the group of GBS and/or GAS and/or *S pneumoniae* serotypes.

In vivo efficacy models include but are not limited to: (i) A murine infection model using human GBS and/or GAS and/or S pneumoniae serotypes; (ii) a murine disease model which is a murine model using a mouse-adapted GBS and/or GAS and/or S pneumoniae strain, such as those

strains outlined above which is particularly virulent in mice and (iii) a primate model using human GBS or GAS or S pneumoniae isolates.

The immune response may be one or both of a TH1 immune response and a TH2 response.

The immune response may be an improved or an enhanced or an altered immune response.

The immune response may be one or both of a systemic and a mucosal immune response.

Preferably the immune response is an enhanced system and/or mucosal response.

An enhanced systemic and/or mucosal immunity is reflected in an enhanced TH1 and/or TH2 immune response. Preferably, the enhanced immune response includes an increase in the production of IgG1 and/or IgG2a and/or IgA

Preferably the mucosal immune response is a TH2 immune response. Preferably, the mucosal immune response includes an increase in the production of IgA.

Activated TH2 cells enhance antibody production and are therefore of value in responding to extracellular infections. Activated TH2 cells may secrete one or more of IL-4, IL-5, IL-6, and IL-10. A TH2 immune response may result in the production of IgG1, IgE, IgA and memory B cells for future protection.

A TH2 immune response may include one or more of an increase in one or more of the cytokines associated with a TH2 immune response (such as IL-4, IL-5, IL-6 and IL-10), or an increase in the production of IgG1, IgE, IgA and memory B cells. Preferably, the enhanced TH2 immune resonse will include an increase in IgG1 production.

A TH1 immune response may include one or more of an increase in CTLs, an increase in one or more of the cytokines associated with a TH1 immune response (such as IL-2, IFNγ, and TNFβ), an increase in activated macrophages, an increase in NK activity, or an increase in the production of IgG2a. Preferably, the enhanced TH1 immune response will include an increase in IgG2a production.

Immunogenic compositions of the invention, in particular, immunogenic composition comprising one or more GAS antigens of the present invention may be used either alone or in combination with other GAS antigens optionally with an immunoregulatory agent capable of eliciting a Th1 and/or Th2 response.

Compositions of the invention will generally be administered directly to a patient. Certain routes may be favored for certain compositons, as resulting in the generation of a more effective immune response, preferably a CMI response, or as being less likely to induce side effects, or as being easier for administration. Direct delivery may be accomplished by parenteral injection (*e.g.* subcutaneously, intraperitoneally, intradermally, intravenously, intramuscularly, or to the interstitial space of a tissue), or by rectal, oral (*e.g.* tablet, spray), vaginal, topical, transdermal (*e.g.* see WO 99/27961) or transcutaneous (*e.g.* see WO 02/074244 and WO 02/064162), intranasal (*e.g.* see WO03/028760), ocular, aural, pulmonary or other mucosal administration.

The invention may be used to elicit systemic and/or mucosal immunity.

In one particularly preferred embodiment, the immunogenic composition comprises one or more GBS or GAS or S pneumoniae antigen(s) which elicits a neutralising antibody response and one or more GBS or GAS or S pneumoniae antigen(s) which elicit a cell mediated immune response. In this way, the neutralising antibody response prevents or inhibits an initial GBS or GAS or S pneumoniae infection while the cell-mediated immune response capable of eliciting an enhanced Th1 cellular response prevents further spreading of the GBS or GAS or S pneumoniae infection. Preferably, the immunogenic composition comprises one or more GBS or GAS or S pneumoniae surface antigens and one or more GBS or GAS or S pneumoniae cytoplasmic antigens. Preferably the immunogenic composition comprises one or more GBS or GAS or S pneumoniae surface antigens or the like and one or other antigens, such as a cytoplasmic antigen capable of eliciting a Th1 cellular response.

Dosage treatment can be a single dose schedule or a multiple dose schedule. Multiple doses may be used in a primary immunisation schedule and/or in a booster immunisation schedule. In a multiple dose schedule the various doses may be given by the same or different routes *e.g.* a parenteral prime and mucosal boost, a mucosal prime and parenteral boost, *etc.*

The compositions of the invention may be prepared in various forms. For example, the compositions may be prepared as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared (*e.g.* a lyophilised composition). The composition may be prepared for topical administration *e.g.* as an ointment, cream or powder. The composition may be prepared for oral administration *e.g.* as a tablet or capsule, as a spray, or as a syrup (optionally flavoured). The composition may be prepared for pulmonary administration *e.g.* as an inhaler, using a fine powder or a spray. The composition may be prepared as a suppository or pessary. The composition may be prepared for nasal, aural or ocular administration *e.g.* as drops. The composition may be in kit form, designed such that a combined composition is reconstituted just prior to administration to a patient. Such kits may comprise one or more antigens in liquid form and one or more lyophilised antigens.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of antigen(s), as well as any other components, such as antibiotics, as needed. By 'immunologically effective amount', it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention, or increases a measurable immune response or prevents or reduces a clinical symptom. This amount varies depending upon the health and physical condition of the individual to be treated, age, the taxonomic group of individual to be treated (*e.g.* non-human primate, primate, *etc.*), the capacity of the individual's immune system to synthesise antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Further Components of the Composition

The composition of the invention will typically, in addition to the components mentioned above, comprise one or more 'pharmaceutically acceptable carriers', which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. The vaccines may also contain diluents, such as water, saline, glycerol, *etc.* Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present. A thorough discussion of pharmaceutically acceptable excipients is available in Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th ed., ISBN: 0683306472.

Adjuvants

Vaccines of the invention may be administered in conjunction with other immunoregulatory agents. In particular, compositions will usually include an adjuvant. Adjuvants for use with the invention include, but are not limited to, one or more of the following set forth below:

A. Mineral Containing Compositions

Mineral containing compositions suitable for use as adjuvants in the invention include mineral salts, such as aluminum salts and calcium salts. The invention includes mineral salts such as hydroxides (*e.g.* oxyhydroxides), phosphates (*e.g.* hydroxyphosphates, orthophosphates), sulfates, *etc.* (*e.g.* see chapters 8 & 9 of *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.), or mixtures of different mineral compounds (*e.g.* a mixture of a phosphate and a hydroxide adjuvant, optionally with an excess of the phosphate), with the compounds taking any suitable form (*e.g.* gel, crystalline, amorphous, *etc.*), and with adsorption to the salt(s) being preferred. The mineral containing compositions may also be formulated as a particle of metal salt (WO 00/23105).

Aluminum salts may be included in vaccines of the invention such that the dose of Al^{3+} is between 0.2 and 1.0 mg per dose.

B. Oil-Emulsions

Oil-emulsion compositions suitable for use as adjuvants in the invention include squalene-water emulsions, such as MF59 (5% Squalene, 0.5% Tween 80, and 0.5% Span 85, formulated into submicron particles using a microfluidizer). See WO90/14837. See also, Podda, "The adjuvanted influenza vaccines with novel adjuvants: experience with the MF59-adjuvanted vaccine", Vaccine (2001) 19: 2673-2680; Frey et al., "Comparison of the safety, tolerability, and immunogenicity of a MF59-adjuvanted influenza vaccine and a non-adjuvanted influenza vaccine in non-elderly adults", Vaccine (2003) 21:4234-4237. MF59 is used as the adjuvant in the FLUAD™ influenza virus trivalent subunit vaccine.

Particularly preferred adjuvants for use in the compositions are submicron oil-in-water emulsions. Preferred submicron oil-in-water emulsions for use herein are squalene/water emulsions optionally containing varying amounts of MTP-PE, such as a submicron oil-in-water emulsion containing 4-5% w/v squalene, 0.25-1.0% w/v Tween 80 ™ (polyoxyelthylenesorbitan monooleate), and/or 0.25-1.0% Span 85™ (sorbitan trioleate), and, optionally, N-acetylmuramyl-L-alanyl-D-isogluatminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-huydroxyphosphophoryloxy)-ethylamine (MTP-PE), for example, the submicron oil-in-water emulsion known as "MF59" (International Publication No. WO 90/14837; US Patent Nos. 6,299,884 and 6,451,325, incorporated herein by reference in their entireties; and Ott et al., "MF59 -- Design and Evaluation of a Safe and Potent Adjuvant for Human Vaccines" in *Vaccine Design: The Subunit and Adjuvant Approach* (Powell, M.F. and Newman, M.J. eds.) Plenum Press, New York, 1995, pp. 277-296). MF59 contains 4-5% w/v Squalene (e.g. 4.3%), 0.25-0.5% w/v Tween 80™, and 0.5% w/v Span 85™ and optionally contains various amounts of MTP-PE, formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA). For example, MTP-PE may be present in an amount of about 0-500 μg/dose, more preferably 0-250 μg/dose and most preferably, 0-100 μg/dose. As used herein, the term "MF59-0" refers to the above submicron oil-in-water emulsion lacking MTP-PE, while the term MF59-MTP denotes a formulation that contains MTP-PE. For instance, "MF59-100" contains 100 μg MTP-PE per dose, and so on. MF69, another submicron oil-in-water emulsion for use herein, contains 4.3% w/v squalene, 0.25% w/v Tween 80™, and 0.75% w/v Span 85™ and optionally MTP-PE. Yet another submicron oil-in-water emulsion is MF75, also known as SAF, containing 10% squalene, 0.4% Tween 80™, 5% pluronic-blocked polymer L121, and thr-MDP, also microfluidized into a submicron emulsion. MF75-MTP denotes an MF75 formulation that includes MTP, such as from 100-400 μg MTP-PE per dose.

Submicron oil-in-water emulsions, methods of making the same and immunostimulating agents, such as muramyl peptides, for use in the compositions, are described in detail in International Publication No. WO 90/14837 and US Patent Nos. 6,299,884 and 6,45 1,325, incorporated herein by reference in their entireties.

Complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA) may also be used as adjuvants in the invention.

C. Saponin Formulations

Saponin formulations, may also be used as adjuvants in the invention. Saponins are a heterologous group of sterol glycosides and triterpenoid glycosides that are found in the bark, leaves, stems, roots and even flowers of a wide range of plant species. Saponin from the bark of the *Quillaia saponaria* Molina tree have been widely studied as adjuvants. Saponin can also be commercially obtained from *Smilax ornata* (sarsaprilla), *Gypsophilla paniculata* (brides veil), and *Saponaria officianalis* (soap root). Saponin adjuvant formulations include purified formulations, such as QS21, as well as lipid formulations, such as ISCOMs.

Saponin compositions have been purified using High Performance Thin Layer Chromatography (HP-LC) and Reversed Phase High Performance Liquid Chromatography (RP-HPLC). Specific purified fractions using these techniques have been identified, including QS7, QS17, QS18, QS21, QH-A, QH-B and QH-C. Preferably, the saponin is QS21. A method of production of QS21 is disclosed in US Patent No. 5,057,540. Saponin formulations may also comprise a sterol, such as cholesterol (see WO96/33739).

Combinations of saponins and cholesterols can be used to form unique particles called Immunostimulating Complexs (ISCOMs). ISCOMs typically also include a phospholipid such as phosphatidylethanolamine or phosphatidylcholine. Any known saponin can be used in ISCOMs. Preferably, the ISCOM includes one or more of Quil A, QHA and QHC. ISCOMs are further described in EP0109942, WO 96/11711 and WO 96/33739. Optionally, the ISCOMS may be devoid of additional detergent. See WO 00/07621.

A review of the development of saponin based adjuvants can be found at Barr, et al., "ISCOMs and other saponin based adjuvants", Advanced Drug Delivery Reviews (1998) 32:247-271. See also Sjolander, et al., "Uptake and adjuvant activity of orally delivered saponin and ISCOM vaccines", Advanced Drug Delivery Reviews (1998) 32:321-338.

D. *Virosomes and Virus Like Particles (VLPs)*

Virosomes and Virus Like Particles (VLPs) can also be used as adjuvants in the invention. These structures generally contain one or more proteins from a virus optionally combined or formulated with a phospholipid. They are generally non-pathogenic, non-replicating and generally do not contain any of the native viral genome. The viral proteins may be recombinantly produced or isolated from whole viruses. These viral proteins suitable for use in virosomes or VLPs include proteins derived from influenza virus (such as HA or NA), Hepatitis B virus (such as core or capsid proteins), Hepatitis E virus, measles virus, Sindbis virus, Rotavirus, Foot-and-Mouth Disease virus, Retrovirus, Norwalk virus, human Papilloma virus, HIV, RNA-phages, Qß-phage (such as coat proteins), GA-phage, fr-phage, AP205 phage, and Ty (such as retrotransposon Ty protein p1). VLPs are discussed further in WO 03/024480, WO 03/024481, and Niikura et al., "Chimeric Recombinant Hepatitis E Virus-Like Particles as an Oral Vaccine Vehicle Presenting Foreign Epitopes", Virology (2002) 293:273-280; Lenz et al., "Papillomarivurs-Like Particles Induce Acute Activation of Dendritic Cells", Journal of Immunology (2001) 5246-5355; Pinto, et al., "Cellular Immune Responses to Human Papillomavirus (HPV)-16 L1 Healthy Volunteers Immunized with Recombinant HPV-16 L1 Virus-Like Particles", Journal of Infectious Diseases (2003) 188:327-338; and Gerber et al., "Human Papillomavrisu Virus-Like Particles Are Efficient Oral Immunogens when Coadministered with Escherichia coli Heat-Labile Entertoxin Mutant R192G or CpG", Journal of Virology (2001) 75(10):4752-4760. Virosomes are discussed further in, for example, Gluck et al., "New Technology Platforms in the Development of Vaccines for the Future", Vaccine (2002) 20:B10 –B16. Immunopotentiating reconstituted influenza virosomes (IRIV) are used as the subunit antigen

delivery system in the intranasal trivalent INFLEXAL™ product {Mischler & Metcalfe (2002) *Vaccine* 20 Suppl 5:B17-23} and the INFLUVAC PLUS™ product.

E. Bacterial or Microbial Derivatives

Adjuvants suitable for use in the invention include bacterial or microbial derivatives such as:

(1) *Non-toxic derivatives of enterobacterial lipopolysaccharide (LPS)*

Such derivatives include Monophosphoryl lipid A (MPL) and 3-O-deacylated MPL (3dMPL). 3dMPL is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains. A preferred "small particle" form of 3 De-O-acylated monophosphoryl lipid A is disclosed in EP 0 689 454. Such "small particles" of 3dMPL are small enough to be sterile filtered through a 0.22 micron membrane (see EP 0 689 454). Other non-toxic LPS derivatives include monophosphoryl lipid A mimics, such as aminoalkyl glucosaminide phosphate derivatives *e.g.* RC-529. See Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278.

(2) *Lipid A Derivatives*

Lipid A derivatives include derivatives of lipid A from *Escherichia coli* such as OM-174. OM-174 is described for example in Meraldi et al., "OM-174, a New Adjuvant with a Potential for Human Use, Induces a Protective Response with Administered with the Synthetic C-Terminal Fragment 242-310 from the circumsporozoite protein of Plasmodium berghei", Vaccine (2003) 21:2485-2491; and Pajak, et al., "The Adjuvant OM-174 induces both the migration and maturation of murine dendritic cells in vivo", Vaccine (2003) 21:836-842.

(3) *Immunostimulatory oligonucleotides*

Immunostimulatory oligonucleotides suitable for use as adjuvants in the invention include nucleotide sequences containing a CpG motif (a sequence containing an unmethylated cytosine followed by guanosine and linked by a phosphate bond). Bacterial double stranded RNA or oligonucleotides containing palindromic or poly(dG) sequences have also been shown to be immunostimulatory.

The CpG's can include nucleotide modifications/analogs such as phosphorothioate modifications and can be double-stranded or single-stranded. Optionally, the guanosine may be replaced with an analog such as 2'-deoxy-7-deazaguanosine. See Kandimalla, et al., "Divergent synthetic nucleotide motif recognition pattern: design and development of potent immunomodulatory oligodeoxyribonucleotide agents with distinct cytokine induction profiles", Nucleic Acids Research (2003) 31(9): 2393-2400; WO02/26757 and WO99/62923 for examples of possible analog substitutions. The adjuvant effect of CpG oligonucleotides is further discussed in Krieg, "CpG motifs: the active ingredient in bacterial extracts?", Nature Medicine (2003) 9(7): 831-835; McCluskie, et al., "Parenteral and mucosal prime-boost immunization strategies in mice with hepatitis B surface antigen and CpG DNA", FEMS Immunology and Medical Microbiology (2002) 32:179-185; WO98/40100; US Patent No. 6,207,646; US Patent No. 6,239,116 and US Patent No. 6,429,199.

The CpG sequence may be directed to TLR9, such as the motif GTCGTT or TTCGTT. See Kandimalla, et al., "Toll-like receptor 9: modulation of recognition and cytokine induction by novel

synthetic CpG DNAs", Biochemical Society Transactions (2003) 31 (part 3): 654-658. The CpG sequence may be specific for inducing a Th1 immune response, such as a CpG-A ODN, or it may be more specific for inducing a B cell response, such a CpG-B ODN. CpG-A and CpG-B ODNs are discussed in Blackwell, et al., "CpG-A-Induced Monocyte IFN-gamma-Inducible Protein-10 Production is Regulated by Plasmacytoid Dendritic Cell Derived IFN-alpha", J. Immunol. (2003) 170(8):4061-4068; Krieg, "From A to Z on CpG", TRENDS in Immunology (2002) 23(2): 64-65 and WO01/95935. Preferably, the CpG is a CpG-A ODN.

Preferably, the CpG oligonucleotide is constructed so that the 5' end is accessible for receptor recognition. Optionally, two CpG oligonucleotide sequences may be attached at their 3' ends to form "immunomers". See, for example, Kandimalla, et al., "Secondary structures in CpG oligonucleotides affect immunostimulatory activity", BBRC (2003) 306:948-953; Kandimalla, et al., "Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic GpG DNAs", Biochemical Society Transactions (2003) 31(part 3):664-658; Bhagat et al., "CpG penta- and hexadeoxyribonucleotides as potent immunomodulatory agents" BBRC (2003) 300:853-861 and WO 03/035836.

(4) *ADP-ribosylating toxins and detoxified derivatives thereof.*

Bacterial ADP-ribosylating toxins and detoxified derivatives thereof may be used as adjuvants in the invention. Preferably, the protein is derived from *E. coli* (i.e., E. coli heat labile enterotoxin "LT), cholera ("CT"), or pertussis ("PT"). The use of detoxified ADP-ribosylating toxins as mucosal adjuvants is described in WO95/17211 and as parenteral adjuvants in WO98/42375. Preferably, the adjuvant is a detoxified LT mutant such as LT-K63, LT-R72, and LTR192G. The use of ADP-ribosylating toxins and detoxified derivaties thereof, particularly LT-K63 and LT-R72, as adjuvants can be found in the following references, each of which is specifically incorporated by reference herein in their entirety: Beignon, et al., "The LTR72 Mutant of Heat-Labile Enterotoxin of Escherichia coli Enahnces the Ability of Peptide Antigens to Elicit CD4+ T Cells and Secrete Gamma Interferon after Coapplication onto Bare Skin", Infection and Immunity (2002) 70(6):3012-3019; Pizza, et al., "Mucosal vaccines: non toxic derivatives of LT and CT as mucosal adjuvants", Vaccine (2001) 19:2534-2541; Pizza, et al., "LTK63 and LTR72, two mucosal adjuvants ready for clinical trials" Int. J. Med. Microbiol (2000) 290(4-5):455-461; Scharton-Kersten et al., "Transcutaneous Immunization with Bacterial ADP-Ribosylating Exotoxins, Subunits and Unrelated Adjuvants", Infection and Immunity (2000) 68(9):5306-5313; Ryan et al., "Mutants of Escherichia coli Heat-Labile Toxin Act as Effective Mucosal Adjuvants for Nasal Delivery of an Acellular Pertussis Vaccine: Differential Effects of the Nontoxic AB Complex and Enzyme Activity on Th1 and Th2 Cells" Infection and Immunity (1999) 67(12):6270-6280; Partidos et al., "Heat-labile enterotoxin of Escherichia coli and its site-directed mutant LTK63 enhance the proliferative and cytotoxic T-cell responses to intranasally co-immunized synthetic peptides", Immunol. Lett. (1999) 67(3):209-216; Peppoloni et al., "Mutants of the Escherichia coli heat-labile enterotoxin as safe and strong adjuvants for intranasal delivery of vaccines", Vaccines (2003) 2(2):285-293; and Pine et al., (2002) "Intranasal

immunization with influenza vaccine and a detoxified mutant of heat labile enterotoxin from Escherichia coli (LTK63)" J. Control Release (2002) 85(1-3):263-270. Numerical reference for amino acid substitutions is preferably based on the alignments of the A and B subunits of ADP-ribosylating toxins set forth in Domenighini et al., Mol. Microbiol (1995) 15(6):1165-1167, specifically incorporated herein by reference in its entirety.

F. Bioadhesives and Mucoadhesives

Bioadhesives and mucoadhesives may also be used as adjuvants in the invention. Suitable bioadhesives include esterified hyaluronic acid microspheres (Singh *et al.* (2001) *J. Cont. Rele.* 70:267-276) or mucoadhesives such as cross-linked derivatives of poly(acrylic acid), polyvinyl alcohol, polyvinyl pyrollidone, polysaccharides and carboxymethylcellulose. Chitosan and derivatives thereof may also be used as adjuvants in the invention. E.g. WO99/27960.

G. Microparticles

Microparticles may also be used as adjuvants in the invention. Microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, *etc.*), with poly(lactide-co-glycolide) are preferred, optionally treated to have a negatively-charged surface (*e.g.* with SDS) or a positively-charged surface (*e.g.* with a cationic detergent, such as CTAB).

H. Liposomes

Examples of liposome formulations suitable for use as adjuvants are described in US Patent No. 6,090,406, US Patent No. 5,916,588, and EP 0 626 169.

I. Polyoxyethylene ether and Polyoxyethylene Ester Formulations

Adjuvants suitable for use in the invention include polyoxyethylene ethers and polyoxyethylene esters. WO99/52549. Such formulations further include polyoxyethylene sorbitan ester surfactants in combination with an octoxynol (WO01/21207) as well as polyoxyethylene alkyl ethers or ester surfactants in combination with at least one additional non-ionic surfactant such as an octoxynol (WO 01/21152).

Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether (laureth 9), polyoxyethylene-9-steoryl ether, polyoxytheylene-8-steoryl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

J. Polyphosphazene (PCPP)

PCPP formulations are described, for example, in Andrianov et al., "Preparation of hydrogel microspheres by coacervation of aqueous polyphophazene solutions", Biomaterials (1998) 19(1-3):109-115 and Payne et al., "Protein Release from Polyphosphazene Matrices", Adv. Drug. Delivery Review (1998) 31(3):185-196.

K. Muramyl peptides

Examples of muramyl peptides suitable for use as adjuvants in the invention include N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-l-alanyl-d-isoglutamine (nor-MDP), and N-acetylmuramyl-l-alanyl-d-isoglutaminyl-l-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE).

L. Imidazoquinolone Compounds.

Examples of imidazoquinolone compounds suitable for use adjuvants in the invention include Imiquamod and its homologues, described further in Stanley, "Imiquimod and the imidazoquinolones: mechanism of action and therapeutic potential" Clin Exp Dermatol (2002) 27(7):571-577 and Jones, "Resiquimod 3M", Curr Opin Investig Drugs (2003) 4(2):214-218.

The invention may also comprise combinations of aspects of one or more of the adjuvants identified above. For example, the following adjuvant compositions may be used in the invention:

(1) a saponin and an oil-in-water emulsion (WO 99/11241);

(2) a saponin (e.g.., QS21) + a non-toxic LPS derivative (e.g. 3dMPL) (see WO 94/00153);

(3) a saponin (e.g.., QS21) + a non-toxic LPS derivative (e.g. 3dMPL) + a cholesterol;

(4) a saponin (*e.g.* QS21) + 3dMPL + IL-12 (optionally + a sterol) (WO 98/57659);

(5) combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions (See European patent applications 0835318, 0735898 and 0761231);

(6) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion.

(7) Ribi™ adjuvant system (RAS), (Ribi Immunochem) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™);

(8) one or more mineral salts (such as an aluminum salt) + a non-toxic derivative of LPS (such as 3dPML).

(9) one or more mineral salts (such as an aluminum salt) + an immunostimulatory oligonucleotide (such as a nucleotide sequence including a CpG motif). Combination No. (9) is a preferred adjuvant combination.

M. Human Immunomodulators

Human immunomodulators suitable for use as adjuvants in the invention include cytokines, such as interleukins (*e.g.* IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (*e.g.* interferon-γ), macrophage colony stimulating factor, and tumor necrosis factor.

Aluminum salts and MF59 are preferred adjuvants for use with injectable influenza vaccines. Bacterial toxins and bioadhesives are preferred adjuvants for use with mucosally-delivered vaccines, such as nasal vaccines.

The immunogenic compositions of the present invention may be administed in combination with an antibiotic treatment regime. In one embodiment, the antibiotic is administered prior to administration of the antigen of the invention or the composition comprising the one or more of the antigens of the invention.

In another embodiment, the antibiotic is administered subsequent to the adminstration of the one or more antigens of the invention or the composition comprising the one or more antigens of the invention. Examples of antibiotics suitable for use in the treatment of the Steptococcal infections of the invention include but are not limited to penicillin or a derivative thereof or clindamycin or the like.

Further antigens

The compositions of the invention may further comprise one or more additional Gram positive bacterial antigens which are not associated with an AI. Preferably, the Gram positive bacterial antigens that are not associated with an AI can provide protection across more than one serotype or strain isolate. For example, a first non-AI antigen, in which the first non-AI antigen is at least 90% (*i.e.*, at least 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100%) homologous to the amino acid sequence of a second non-AI antigen, wherein the first and the second non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria, may be further included in the compositions. The first non-AI antigen may also be homologous to the amino acid sequence of a third non-AI antigen, such that the first non-AI antigen, the second non-AI antigen, and the third non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria. The first non-AI antigen may also be homologous to the amino acid sequence of a fourth non-AI antigen, such that the first non-AI antigen, the second non-AI antigen, the third non-AI antigen, and the fourth non-AI antigen are derived from the genomes of different serotypes of a Gram positive bacteria.

The first non-AI antigen may be GBS 322. The amino acid sequence of GBS 322 across GBS strains from serotypes Ia, Ib, II, III, V, and VIII is greater than 90%. Alternatively, the first non-AI antigen may be GBS 276. The amino acid sequence of GBS 276 across GBS strain from serotypes Ia, Ib, II, III, V, and VIII is greater than 90%. Table 13 provides the percent amino acid sequence identity of GBS 322 and GBS 276 across different GBS strains and serotypes.

Table 13: Conservation of GBS 322 and GBS 276 amino acid sequences

Serotype	**Strains**	**GBS 322**		**GBS 276**	
		cGH	%AA identity	cGH	%AA identity
Ia	**090**	+	98.60	+	97.90
	A909	+	98.30	+	97.90
	515	+	98.80	+	97.50
	DK1	+		+	
	DK8	+		+	
	Davis	+		+	
Ib	**7357b**	+		+	
	H36B	+	98.30	+	97.80
II	**18RS21**	+	100.00	+	99.90
	DK21	+		+	

Serotype	Strains	GBS 322		GBS 276	
		cGH	%AA identity	cGH	%AA identity
III	NEM316	+	100.00	+	97.00
	COH31	+		+	
	D136	+		+	
	M732	+	98.00	+	100.00
	COH1	+	98.30	+	100.00
	M781	+	98.30	+	99.60
No type	CJB110	+	98.60	+	97.90
	1169NT	+	97.40	+	97.90
V	CJB111	+	100.00	+	
	2603	+	100.00	+	100.00
VIII	JM130013	+	100.00	+	97.90
	SMU014	+		+	
total		22/22	98.28+/-0.4	22/22	98.44 +/-1.094

As an example, inclusion of a non-AI protein, GBS 322, in combination with AI antigens GBS 67, GBS 80, and GBS 104 provided protection to newborn mice in an active maternal immunization assay.

Table 14: Active maternal immunization assay for a combination of fragments from GBS 322, GBS 80, GBS 104, and GBS 67

		FACS (Δ Mean)			MIX=322+80+104+67		PBS	
GBS strains	Type	GBS 80	GBS 67	GBS 322	alive/treated	% protection	alive/treated	% protection
515	Ia	0	409	227	39/40	97	6/40	15
7357b-	Ib	91	316	102	19/30	63	1/30	3
DK21	II	0	331	416	25/34	73	17/48	35
5401	II	170	618	135	35/40	87	3/37	8
3050	II	43	460	188	48/48	100	1/30	3
COH1	III	305	0	130	36/36	100	7/40	17
M781	III	65	0	224	30/40	75	4/39	10
2603	V	125	105	313	27/33	82	10/35	28
CJB111	V	370	481	63	25/28	89	4/46	9
JM9130013	VIII	597	83	143	37/39	95	5/40	12
JMU071	VIII	556	79	170	44/50	88	18/50	36
NT1169	NT	0	443	213	12/32	37	11/35	31

In fact, the non-AI GBS 322 antigen may itself provide protection to newborn mice in an active maternal immunization assay.

Table 16: Active maternal immunization assay for each of GBS 80 and GBS 322 antigens

		GBS 80			GBS 322		
		FACS	Protection (% survival)		FACS	Protection (% survival)	
GBS strains	Type	Δ Mean	antigen	ctrl-	Δ Mean	antigen	ctrl-
CJB111	V	370	72 %	40%	63	57%	40%
COH1	III	305	76 %	10%	130	3%	10%
2603	V	82	22 %	34%	313	83 %	34%
7357b-	Ib	91	36%	34%	102	43%	34%
18RS21	II	0	15%	24%	268	84 %	24%
DK21	II	0	10%	21%	416	67 %	25%
A909	Ia	0	0%	14%			
090	Ia	0	0%	0%			
H36B	Ib				105	34%	32%

Thus, inclusion of a non-AI protein in an immunogenic composition of the invention may provide increased protection a mammal.

The immunogenic compositions comprising *S. pneumonaie* AI polypeptides may further secondary SP protein antigens which include (a) any of the SP protein antigens disclosed in WO 02/077021 or U.S. provisional application __________, filed April 20, 2005 (Attorney Docket Number 002441.00154), (2) immunogenic portions of the antigens comprising at least 7 contiguous amino acids, (3) proteins comprising amino acid sequences which retain immunogenicity and which are at least 95% identical to these SP protein antigens (*e.g.*, 95%, 96%, 97%, 98%, 99%, or 99.5% identical), and (4) fusion proteins, including hybrid SP protein antigens, comprising (1)-(3).

Alternatively, the invention may include an immunogenic composition comprising a first and a second Gram positive bacteria non-AI protein, wherein the polynucleotide sequence encoding the sequence of the first non-AI protein is less than 90% (i.e., less than 90, 88, 86, 84, 82, 81, 78, 76, 74, 72, 70, 65, 60, 55, 50, 45, 40, 35, or 30 percent) homologous than the corresponding sequence in the genome of the second non-AI protein.

The compositions of the invention may further comprise one or more additional non-Gram positive bacterial antigens, including additional bacterial, viral or parasitic antigens. The compositions of the invention may further comprise one or more additional non-GBS antigens, including additional bacterial, viral or parasitic antigens.

In another embodiment, the GBS antigen combinations of the invention are combined with one or more additional, non-GBS antigens suitable for use in a vaccine designed to protect elderly or immunocomprised individuals. For example, the GBS antigen combinations may be combined with an antigen derived from the group consisting of *Enterococcus faecalis, Staphylococcus aureus, Staphylococcus epidermis, Pseudomonas aeruginosa, Legionella pneumophila, Listeria monocytogenes, Neisseria meningitides,* influenza, and Parainfluenza virus ('PIV').

Where a saccharide or carbohydrate antigen is used, it is preferably conjugated to a carrier protein in order to enhance immunogenicity {*e.g.* Ramsay *et al.* (2001) *Lancet* 357(9251):195-196; Lindberg (1999) *Vaccine* 17 Suppl 2:S28-36; Buttery & Moxon (2000) *J R Coll Physicians Lond* 34:163-168; Ahmad & Chapnick (1999) *Infect Dis Clin North Am* 13:113-133, vii.; Goldblatt (1998) *J. Med. Microbiol.* 47:563-567; European patent 0 477 508; US Patent No. 5,306,492; International patent application WO98/42721; *Conjugate Vaccines* (eds. Cruse *et al.*) ISBN 3805549326, particularly vol. 10:48-114; and Hermanson (1996) *Bioconjugate Techniques* ISBN: 0123423368 or 012342335X}. Preferred carrier proteins are bacterial toxins or toxoids, such as diphtheria or tetanus toxoids. The CRM_{197} diphtheria toxoid is particularly preferred {*Research Disclosure*, 453077 (Jan 2002)}. Other carrier polypeptides include the *N.meningitidis* outer membrane protein (EP-A-0372501), synthetic peptides (EP-A-0378881; EP-A-0427347), heat shock proteins (WO 93/17712; WO 94/03208), pertussis proteins (WO 98/58668; EP A 0471177), protein D from *H.influenzae* (WO 00/56360), cytokines (WO 91/01146), lymphokines, hormones, growth factors, toxin A or B from *C.difficile* (WO00/61761), iron-uptake proteins (WO01/72337), *etc.* Where a mixture comprises capsular saccharides from both serogroups A and C, it may be preferred that the ratio (w/w) of MenA saccharide:MenC saccharide is greater than 1 (*e.g.* 2:1, 3:1, 4:1, 5:1, 10:1 or higher). Different saccharides can be conjugated to the same or different type of carrier protein. Any suitable conjugation reaction can be used, with any suitable linker where necessary.

Toxic protein antigens may be detoxified where necessary *e.g.* detoxification of pertussis toxin by chemical and/or genetic means.

Where a diphtheria antigen is included in the composition it is preferred also to include tetanus antigen and pertussis antigens. Similarly, where a tetanus antigen is included it is preferred also to include diphtheria and pertussis antigens. Similarly, where a pertussis antigen is included it is preferred also to include diphtheria and tetanus antigens.

Antigens in the composition will typically be present at a concentration of at least 1μg/ml each. In general, the concentration of any given antigen will be sufficient to elicit an immune response against that antigen.

As an alternative to using protein antigens in the composition of the invention, nucleic acid encoding the antigen may be used {*e.g.* refs. Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648; Scott-Taylor & Dalgleish (2000) *Expert Opin Investig Drugs* 9:471-480; Apostolopoulos & Plebanski (2000) *Curr Opin Mol Ther* 2:441-447; Ilan (1999) *Curr Opin Mol Ther* 1:116-120; Dubensky *et al.* (2000) *Mol Med* 6:723-732; Robinson & Pertmer (2000) *Adv Virus Res* 55:1-74; Donnelly *et al.* (2000) *Am J Respir Crit Care Med* 162(4 Pt 2):S190-193; and Davis (1999) *Mt. Sinai J. Med.* 66:84-90}. Protein components of the compositions of the invention may thus be replaced by nucleic acid (preferably DNA *e.g.* in the form of a plasmid) that encodes the protein.

Definitions

The term "comprising" means "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

The term "about" in relation to a numerical value *x* means, for example, $x \pm 10\%$.

References to a percentage sequence identity between two amino acid sequences means that, when aligned, that percentage of amino acids are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in section 7.7.18 of *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.*, eds., 1987) Supplement 30. A preferred alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62. The Smith-Waterman homology search algorithm is disclosed in Smith & Waterman (1981) *Adv. Appl. Math.* 2: 482-489.

The invention is further illustrated, without limitation, by the following examples.

EXAMPLE 1: Binding of an Adhesin Island surface protein, GBS 80, to Fibrinogen and Fibronectin.

This example demonstrates that an Adhesin Island surface protein, GBS 80 can bind to fibrinogen and fibronectin.

An enzyme-linked immunosorbent assay (ELISA) was used to analyse the *in vitro* binding ability of recombinant GBS 80 to immobilized extra-cellular matrix (ECM) proteins but not to bovine serum albumin (BSA). Microtiter plates were coated with ECM proteins (fibrinogen, fibronectin, laminin, collagen type IV) and binding assessed by adding varying concentrations of a recombinant form of GBS 80, over-expressed and purified from *E. coli* (FIGURE 5A). Plates were then incubated sequentially with a) mouse anti-GBS 80 primary antibody; b) rabbit anti-mouse AP-conjugated secondary antibody; c) pNPP colorimetric substrate. Relative binding was measured by monitoring absorbance at 405 nm, using 595 nm as a reference wavelength. Figure 5b shows binding of recombinant GBS 80 to immobilized ECM proteins (1 μg) as a function of concentration of GBS 80. BSA was used as a negative control. Data points represent the means of OD_{405} values ± standard deviation for 3 wells.

Binding of GBS 80 to the tested ECM proteins was found to be concentration dependent and exhibited saturation kinetics. As is also evident from FIGURE 5, binding of GBS 80 to fibronectin and fibrinogen was greater than binding to laminin and collagen type IV at all the concentrations tested.

EXAMPLE 2: GBS 80 is required for surface localization of GBS 104.

This example demonstrates that co-expression of GBS 80 is required for surface localization of GBS 104.

The polycistronic nature of the Adhesin Island I mRNA was investigated through reverse transcriptase-PCR (RT-PCR) analysis employing primers designed to detect transcripts arising from contiguous genes. Total RNA was isolated from GBS cultures grown to an optical density at 600 nm

(OD_{600}) of 0.3 in THB (Todd-Hewitt broth) by the RNeasy Total RNA isolation method (Qiagen) according to the manufacturer's instructions. The absence of contaminating chromosomal DNA was confirmed by failure of the gene amplification reactions to generate a product detectable by agarose gel electrophoresis, in the absence of reverse transcriptase. RT-PCR analysis was performed with the Access RT-PCR system (Promega) according to the manufacturer's instructions, employing PCR cycling temperatures of 60°C for annealing and 70°C for extension. Amplification products were visualized alongside 100-bp DNA markers in 2% agarose gels after ethidium bromide staining.

FIGURE 5 shows that all the genes are co-transcribed as an operon. A schematic of the AI-1 operon is shown above the agarose gel analysis of the RT-PCR products. Large rectangular arrows indicate the predicted transcript direction. Primer pairs were selected such as "1-4" cross the 3'finish-5'start of successive genes and overlap each gene by at least 200 bp. Additionally, "1" crosses a putative rho-independent transcriptional terminator. "5" is an internal GBS 80 control and "6" is an unrelated control from a highly expressed gene. Lanes: "a": RNA plus RTase enzyme; "b" RNA without RTase; "c": genomic DNA control.

In the effort to elucidate the functions of the AI-1 proteins, in frame deletions of all of the genes within the operon have been constructed and the resulting mutants characterized with respect to surface exposure of the encoded antigens (see FIGURE 8).

Each in-frame deletion mutation was constructed by splice overlap extension PCR (SOE-PCR) essentially as decribed by Horton et al. [Horton R. M., Z. L. Cai, S. N. Ho, L. R. Pease (1990) Biotechniques 8:528-35] using suitable primers and cloned into the temperature sensitive shuttle vector pJRS233 to replace the wild type copy by allelic exchange [Perez-Casal, J., J. A. Price, et al. (1993) Mol Microbiol 8(5): 809-19.]. All plasmid constructions utilized standard molecular biology techniques, and the identities of DNA fragments generated by PCR were verified by sequencing. Following SOE-PCR, the resulting mutant DNA fragments were digested with XhoI and EcoRI, and ligated into a similarly digested pJRS233. The resuting vectors were introduced by electroporation into the chromosome of 2603 and COH1 GBS strains in a three-step process, essentially as described in Framson et al. [Framson, P. E., A. Nittayajarn, J. Merry, P. Youngman, and C. E. Rubens. (1997) Appl. Environ. Microbiol. 63(9):3539-47]. Briefly, the vector pJRS233 contains an *erm* gene encoding erythromycin resistance and a temperature-sensitive gram-positive replicon that is active at 30°C but not at 37°C. Initially, the constructs are electroporated into GBS electro-competent cells prepared as described by Frameson et al., and transformants containing free plasmid are selected by their ability to grow at 30°C on Todd-Hewitt Broth (THB) agar plates containing 1 μg/ml erythromycin. The second step includes a selection step for strains in which the plasmid has integrated into the chromosome via a single recombination event over the homologous plasmid insert and chromosome sequence by their ability to grow at 37°C on THB agar medium containing 1 mg/ml erythromycin. In the third step, GBS cells containing the plasmid integrated within the chromosome (integrants) are serially passed in broth culture in the absence of antibiotics at 30°C. Plasmid excision

from the chromosome via a second recombination event over the duplicated target gene sequence either completed the allelic exchange or reconstituted the wild-type genotype. Subsequent loss of the plasmid in the absence of antibiotic selection pressure resulted in an erythromycin-sensitive phenotype. In order to assess gene replacement a screening of erythromycin-sensitive colonies was performed by analysis of the target gene PCR amplicons.

FIGURE 7 reports a schematic of the IS-1 operon for each knock-out strain generated, along with the deletion position within the amino acidic sequence. Most data presented here concern the COH1 deletion strains, in which the expression of each of the antigens is higher by DNA microarray analysis (data not shown) as well as detectable by FACS analysis (see FIGURE 8). The double mutant in 2603 Δ80, Δ104 double mutant was constructed by sequential allelic exchanges of the shown alleles.

Immunization protocol

Immune sera for FACS experiments were obtained as follows.

Groups of 4 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were immunized with the selected GBS antigens, (20 μg of each recombinant GBS antigen), suspended in 100 μl of PBS. Each group received 3 doses at days 0, 21 and 35. Immunization was performed through intra-peritoneal injection of the protein with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. In each immunization scheme negative and positive control groups are used. Immune response was monitored by using serum samples taken on day 0 and 49.

FACS analysis

Preparation of paraformaldehyde treated GBS cells and their FACS analysis were carried out as follows.

GBS serotype COH1 strain cells were grown in Todd Hewitt Broth (THB; Difco Laboratories, Detroit, Mich.) to OD600nm = 0.5. The culture was centrifuged for 20 minutes at 5000 rpm and bacteria were washed once with PBS, resuspended in PBS containing 0.05% paraformaldehyde, and incubated for 1 hours at 37 °C and then overnight at 4°C. 50μl of fixed bacteria (OD600 0.1) were washed once with PBS, resuspended in 20μl of Newborn Calf Serum, (Sigma) and incubated for 20 min. at room temperature. The cells were then incubated for 1 hour at 4°C in 100μl of preimmune or immune sera, diluted 1:200 in dilution buffer (PBS, 20% Newborn Calf Serum, 0.1% BSA). After centrifugation and washing with 200μl of washing buffer (0.1% BSA in PBS), samples were incubated for 1 hour at 4°C with 50μl of R-Phicoerytrin conjugated F(ab)2 goat anti-mouse IgG (Jackson ImmunoResearch Laboratories; Inc.), diluted 1:100 in dilution buffer. Cells were washed with 200μl of washing buffer and resuspended in 200μl of PBS. Samples were analysed using a FACS Calibur apparatus (Becton Dickinson, Mountain View, Calf.) and data were analyzed using the Cell Quest Software (Becton Dickinson). A shift in mean fluorescence intensity of > 75 channels compared to preimmune sera from the same mice was considered positive. This cutoff

was determined from the mean plus two standard deviations of shifts obtained with control sera raised against mock purified recombinant proteins from cultures of *E. coli* carrying the empty expression vector and included in every experiment. Artifacts due to bacterial lysis were excluded using antisera raised against 6 different known cytoplasmic proteins all of which were negative

FACS data on COH1 single KO mutants for GBS 104 and GBS 80 indicated that GBS 80 is required for surface localization of GBS 104.

As shown in FIGURE 8, GBS 104 is not surface exposed in the Δ80 strain (second column, bottom), but is present in the whole protein extracts (see FIGURE 10). Mean shift values suggest that GBS 104 is partially responsible for GBS 80 surface exposure (Mean shift of GBS 80 is reduced to ~60% wild-type levels in Δ104), and that GBS 80 is over-expressed in the complemented strain (mean shift value ~200% wild-type level). The Δ80/pGBS 80 strain contains the GBS 80 orf cloned in the shuttle-vector pAM401 (Wirth, R., F. Y. An, et al. (1986). J Bacteriol 165(3): 831-6). The vector alone does not alter the secretion pattern of GBS 104 (right column). FACS was performed on mid-log fixed bacteria with mouse polyclonal antibodies as indicated at left. Black peak is pre-immune sera, colored peaks are sera from immunized animals.

EXAMPLE 3: Deletion of GBS 80 causes attenuation *in vivo*.

This example demonstrates that deletion of GBS 80 causes attenuation *in vivo*, suggesting that this protein contributes to bacterial virulence.

By using a mouse animal model, we studied the role of GBS 80 and GBS 104 in the virulence of S. agalactiae.

Groups of ten outbred female mice 5-6 week weeks old (Charles River Laboratories, Calco Italy) were inoculated intraperitoneally with different dilutions of the mutant strains and LD50 (lethal dose 50) were calculated according to the method of Reed and Muench [Reed, L. J. and H. Muench (1938).The American Journal of Hygiene 27(3): 493-7]. As presented in the table below the number of colony forming units (cfu) counted for both the Δ80 and the Δ80, Δ104 double mutants is about 10 fold higher when compared to the wild type strain suggesting that inactivation of GBS 80 but not GBS 104 is responsible for an attenuation in virulence. This finding indicates that GBS 80 gene in the AI-1 might contribute to virulence.

Table Lethal dose 50% analysis of AI-1 mutants in the 2603 strain background. LD50s were performed by IP injection of female CD1 mice at an age of 5-6 weeks. LD50s were calculated by the method of Reed and Muench (8).

GBS strain	LD_{50}, cfu	Number of Experiments
Wild Type 2603	2×10^8	4
Δ104 mutant	$\sim 2 \times 10^8$	1
Δ80 mutant	2.6×10^9	3
Δ80, Δ104 double mutant	$\sim 2 \times 10^9$	1

EXAMPLE 4: Effect of Adhesin Island Sortase Deletions on Surface Antigen Presentation

This example demonstrates the effect of adhesin island sortase deletions on surface antigen presentation.

FACS analysis results set forth in FIGURE 9 show that a deletion in sortase SAG0648 prevented GBS 104 from reaching the surface and slightly reduced the surface exposure of GBS 80 (fourth panel; mean shift value ~60% wild-type COH1). In the double sortase knock-out strain, neither antigen was surface exposed (far right panel). Either sortase alone was sufficient for GBS 80 to arrive at the bacterial surface (third and fourth columns, top). No effect was seen on surface exposure of antigens GBS 80 or GBS 104 in the ΔGBS 52 strain. Antibodies derived from purified GBS 52 were either non-specific or were FACS negative for GBS 52 (data not shown). FACS analysis was performed as described above (see EXAMPLE 2).

As shown in FIGURE 10, inactivation of GBS 80 has no effect on GBS 104 expression as much as GBS 104 knock out doesn't change the total amount GBS 80 expressed. The Western blot of whole protein extracts (strains noted above lanes) probed with anti-GBS 80 antisera is shown in panel A. Arrow indicates expected size of GBS 80 (60 kDa). GBS 80 antibodies recognize a doublet, the lower band is not present in ΔGBS 80 strains. Panel B shows a Western blot of whole protein extracts probed with anti-GBS 104 antisera. Arrow indicates expected size of GBS 104 (99.4 kDa). Protein extracts were prepared from the same bacterial cultures used for FACS (FIGURES 8 and 9). In conclusion, although GBS 104 does not arrive at the surface in the Δ80 strain by FACS (FIGURE 8, second column), it is present at approximately wild-type levels in the whole protein preps (B, second lane). Approximately 20 μg of each protein extract was loaded per lane.

Western-blot analysis

Aliquots of total protein extract mixed with SDS loading buffer (1x: 60 mM TRIS-HCl pH 6.8, 5% w/v SDS, 10% v/v glycerin, 0.1% Bromophenol Blue, 100 mM DTT) and boiled 5 minutes at 95° C, were loaded on a 12.5% SDS-PAGE precast gel (Biorad). The gel is run using a SDS-PAGE running buffer containing 250 mM TRIS, 2.5 mM Glycine and 0.1 %SDS. The gel is electroblotted onto nitrocellulose membrane at 200 mA for 60 minutes. The membrane is blocked for 60 minutes with PBS/0.05 % Tween-20 (Sigma), 10% skimmed milk powder and incubated O/N at 4° C with PBS/0.05 % Tween 20, 1% skimmed milk powder, with the appropriate dilution of the sera. After washing twice with PBS/0.05 % Tween, the membrane is incubated for 2 hours with peroxidase-conjugated secondary anti-mouse antibody (Amersham) diluted 1:4000. The nitrocellulose is washed three times for 10 minutes with PBS/0.05 % Tween and once with PBS and thereafter developed by Opti-4CN Substrate Kit (Biorad).

Example 5: Binding of Adhesin Island proteins to epithelial cells and effect of Adhesin Island proteins on capacity of GBS to adhere to epithelial cells.

This example illustrates the binding of AI proteins to epithelial cells and the effect of AI proteins on the capacity of GBS to adhere to epithelial cells.

Applicants analysed whether recombinant AI surface proteins GBS 80 or GBS 104 would demonstrate binding to various epithelial cells in a FACS analysis. Applicants also analysed whether

deletion of AI surface proteins GBS 80 or GBS 104 would effect the capacity of GBS to adhere to and invade ME180 cervical epithelial cells.

As shown in Figure 28, deletion of GBS 80 sequence from GBS strain isolate 2603 (serotype V) did not affect the capacity of the mutated GBS to adhere to and invade ME180 cervical epithelial cells. Here ME180 cervical carcinoma epithelial cells were infected with wild type GBS 2603 or GBS 2603 Δ80 isogenic mutant. After two hours of infection, non-adherent bacteria were washed off and infection prolonged for a further two hours and four hours. In invasion experiments, after each time point, was followed by a two hour antibiotic treatment. Cells were then lysed with 1% saponin and lysates platedon TSA plates. As shown in Figure 28, there was little difference between the percent invasion or percent adhesion of wild type and mutant strains up to the four hour time point.

Figure 30 repeats this experiment with both Δ104 and Δ 80 mutants from a different strain isolate. Here, ME180 cervical carcinoma epithelial cells were infected with GBS strain isolate COH (serotype III) wild type or COH1 ΔGBS 104 or COH1 Δ80 isogenic mutant. After one hour of infection, non-adherent bacteria were washed off and the cells were lysed with 1% saponin. The lysates were plated on TSA plates. As shown in Figure 30, while there was little difference in the percent invasion, there was a significant decrease in the percent association of the Δ104 mutant compared to both the wild type and Δ80 mutant.

The affect of AI surface proteins on the ability of GBS to translocate through an epithelial monolayer was also analysed. As shown in Figure 31, a GBS 80 knockout mutant strain partially loses the ability to translocate through an epithelial monolayer. Here epithelial monolayers were inoculated with wildtype or knockout mutant in the apical chamber of a transwell system for two hours and then non-adherent bacteria were washed off. Infection was prolonged for a further two and four hours. Samples were taken from the media of the basolateral side and the number of colony forming unties measured. Transepithelial electrical resistance measured prior to and after infection gave comparable values, indicating the maintenance of the integrity of the monolayer. By the six hour time point, the Δ80 mutants demonstrated a reduced percent transcytosis.

A similar experiment was conducted with GBS 104 knock out mutants. Here, as shown in Figure 22, the Δ104 mutants also demonstrated a reduced percent transcytosis, indicating that the mutant strains translocate through an epithelial monolayer less efficiently than their isogenic wild type counterparts.

Applicants also studied the effect of AI proteins on the capacity of a GBS strain to invade J774 macrophage-like cells. Here, J774 cells were infected with GBS COH1 wild type or COH1 ΔGBS104 or COH1 ΔGBS80 isogenic mutants. After one hour of infection, non-adherent bacteria were washed off and intracellular bacteria were recovered at two, four and six hours post antibiotic treatment. At each time point, cells were lysed with 0.25% Triton X-100 and lysates plated on TSA plates. As shown in Figure 32, the Δ104 mutant demonstrated a significantly reduced percent invasion compared to both the wild type and Δ80 mutant.

Example 6: Hyperoligomeric structures comprising AI surface proteins GBS 80 and GBS 104.

This example illustrates hyperoligomeric structures comprising AI surface proteins GBS 80 and GBS 104. A GBS isolate COH1 (serotype III) was adapted to increase expression of GBS 80. Figure 34 presents a regular negative stain electron micrograph of this mutant; no pilus or hyperoligomeric structures are distinguishable on the surface of the bacteria. When the EM stain is based on anti-GBS 80 antibodies labelled with 10 or 20 nm gold particles, the presence of GBS 80 throughout the hyperoligomeric structure is clearly indicated (Figures 36, 37 and 38). EM staining against GBS 104 (anti-GBS 104 antibodies labelled with 10 nm gold particles) also reveals the presence of GBS 104 primarily on or near the surface of the bacteria or potentially associated with bacterial peptidoglycans (Figure 39). Analysis of this same strain (over-expressing GBS 80) with a combination of both anti-GBS 80 (using 20 nm gold particles) and anti-GBS 104 (using 10 nm gold particles) reveals the presence of GBS 104 on the surface and within the hyperoligomeric structures (see Figures 40 and 41).

Example 7: GBS 80 is necessary for polymer formation and GBS 104 and sortase SAG0648 are necessary for efficient pili assembly

This example demonstrates that GBS 80 is necessary for formation of polymers and that GBS 104 and sortase SAG0648 are necessary for efficient pili assembly. GBS 80 and GBS 104 polymeric assembly was systematically analyzed in Coh1 strain single knock out mutants of each of the relevant coding genes in AI-1 (GBS 80, GBS 104, GBS 52, sag0647, and sag0648). Figure 41 provides Western blots of total protein extracts (strains noted above lanes) probed with either anti-GBS 80 (left panel) sera or anti-GBS 104 sera (right panel) for each of these Coh1 and Coh1 knock out strains. (Coh1, wild type Coh1; Δ80, Coh1 with GBS 80 knocked out; Δ104, Coh1 with GBS 104 knocked out; Δ52, Coh1 with GBS 52 knocked out; Δ647, Coh1 with SAG0647 knocked out; Δ648, Coh1 with SAG0648 knocked out, Δ647-8, Coh1 with SAG0647 and SAG0648 knocked out; Δ80/pGBS80, Coh1 with GBS 80 knocked out but complemented with a high copy number plasmid expressing GBS 80. Asterisks identify the monomer of GBS 80 and GBS 104.)

The smear of immunoreactive material observed in the wild type strain, along with its disappearance in Δ80 and Δ104 mutants, is consistent with the notion that such high molecular weight structures are composed of covalently linked (SDS-resistant) GBS 80 and GBS 104 subunits. The immunoblotting with both anti-GBS 80 (α-GBS 80) and anti-GBS 104 (α-GBS 104) revealed that deletion of sortase SAG0648 also interferes with the assembly of high molecular weight species, whereas the knock out mutant of the second sortase (SAG0647), even if somehow reduced, still maintains the ability to form polymeric structures.

Total extracts form GBS were prepared as follows. Bacteria were grown in 50 ml of Todd-Hewitt broth (Difco) to an OD_{600nm} of 0.5-0.6 and successively pelleted. After two washes in PBS the pellet was resuspended and incubated 3 hours at 37°C with mutanolisin. Cells were then lysed with at

least three freezing-thawing cycles in dry ice and a 37°C bath. The lysate was then centrifuged to

eliminate the cellular debris and the supernatant was quantified. Approximately 40 μg of each protein extract was separated on SDS-PAGE. The gel was then subjected to immunoblotting with mice antisera and detected with chemiluminescence.

Example 8: GBS 80 is polymerized by an AI-2 sortase

This example illustrates that GBS 80 can be polymerized not only by AI-1 sortases, but also by AI-2 sortases. Figure 42 shows total cell extract immunoblots of GBS 515 strain, which lacks AI-1. The left panel, where an anti-GBS 67 sera was used, shows that GBS 67 from AI-2 is assembled into high-molecular weight-complexes, suggesting the formation of a second type of pilus. The same high molecular structure is observed when GBS 80 is highly expressed by reintroducing the gene within a plasmid (pGBS 80). By using anti-GBS 80 (right panel) sera on the same extracts, again it is observed that, with GBS 80 over expression (515/pGBS 80), a high-molecular weight structure is assembled. This implies that, in the absence of AI-1 sortases, AI-2 sortases (SAG1405 and SAG1406) can complement the lacking function, still being able to assemble GBS 80 in a pilus structure.

Example 9: Coh1 produces a high molecular weight molecule, the GBS 80 pilin

This example illustrates that Coh1 produces a high molecular weight molecule, greater than 1000 kDa, which is the GBS 80 pilin. Figure 43 provides silver-stained electrophoretic gels that show that Coh1 produces two macromolecules. One of these macromolecules disappears in the Coh1 GBS 80 knock out cells, but does not disappear in the Coh1 GBS 52 knock out mutant cells. The last two lanes on the right were loaded with 15 times the amount loaded in the other lanes. This was done in order to be able to count the bands. By doing this, a conservative size estimate of the top bands was calculated by starting at 240 kDa and considering each of 14 higher bands as the result of consecutive additions of a GBS 80 monomer.

Coh1, wild type Coh1; Δ80, Coh1 cells with GBS 80 knocked out; Δ52, Coh1 cells with GBS 52 knocked out; Δ80/pGBS 80, Coh1 cells with GBS 80 knocked out and complemented with a high copy number construct expressing GBS 80.

Example 10. GBS 52 is a minor component of the GBS pilus

This example illustrates that GBS 52 is present in the GBS pilus and is a minor component of the pilus. Figure 45 shows an immunoblot of total cell extracts from a GBS Coh1 strain and a GBS Coh1 strain knocked out for GBS 52 (Δ52). The total cell extracts were immunoblotted anti-GBS 80 antisera (left) and anti-GBS 52 antisera (right). Immunoblotting was performed using a 3-8% Tris-acetate polyacrylamide gel (Invitrogen) which provided excellent separation of large molecular weight proteins (see figure 41). When the gel was incubated with anti-GBS 80 sera, the bands from the Coh1 wild-type strain appeared shifted when compared to the Δ52 mutant. This observation

indicated a different size of the pilus polymeric components in the two strains. When the same gel was stripped and incubated with anti-GBS 52 sera the high-molecular subunits in the Coh1 wild-type strain showed similar molecular size of those in the correspondent lane in the left panel. These findings confirmed that GBS 52 is indeed associated with GBS 80 macro-molecular structures but represents a minor component of the GBS pilus.

Example 11: Pilus structures are present in the supernatant of GBS bacterial cultures

This example illustrates that the pilus structure assembled in Coh1 GBS is present in the supernatant of a bacterial cell culture. Figure 46 shows an immunoblot where the protein extract of the supernatant from cultures of different GBS mutant strains (117 = Coh1 GBS 80 knockout; 159= Coh1 GBS 104 knockout; 202= Coh1 GBS 52 knockout; 206= Coh1 GBS sag0647 knockout; 208= Coh1 GBS sag0648 knockout; 197= Coh1 GBS sag0647/sag0648 knockout; 179= Coh1 GBS 80 knockout complemented with a high copy plasmid expressing GBS 80). GBS 80 antisera detects the presence of pilus structures in the appropriate Coh1 strains.

The protein extract was prepared as follows. Bacteria were grown in THB to an OD_{600nm} of 0.5-0.6 and the supernatant was separated from the cells by centrifugation. The supernatant was then filtered (∅ 0.2 μm) and 1 ml was added with 60% TCA for protein precipitation.
GBS pili were also extracted from the fraction of surface-exposed proteins in Coh1 strain and its GBS 80 knock out mutant as described hereafter. Bacteria were grown to an OD_{600nm} of 0.6 in 50 ml of THB at 37°C. Cells were washed once with PBS and the pellet was then resuspended in 0.1 M KPO4 pH 6.2, 40% sucrose, 10 mM MgCl2, 400U/ml mutanolysin and incubated 3 hours at 37°C. Protoplasts were separated by centrifugation and the supernatant was recovered and its protein content measured.

In order to study the dynamics of pilus production during different growth phases, 1 ml supernatant of a culture at different OD_{600nm} was TCA precipitated and loaded onto a 3-8% SDS-PAGE as described before. Figure 47 shows the corresponding Western blot with GBS 80 anti-sera. The first group of lanes (left five sample lanes) refer to a Coh1 strain growth (OD_{600nm} are noted above the lanes) whereas the second group of lanes (right five samples) are from a GBS 80 knock out strain over expressing GBS 80. The experiment shows that pilus macromolecular structures can be found in the supernatant in all of the growth phases tested.

Example 12: In GBS strain Coh1, only GBS 80 and a sortase (sag0647 or sag0648) is required for polymerization

This example describes requirements for pilus formation in Coh1. Figure 48 shows a Western blot of total protein extracts (prepared as described before) using anti-GBS 80 sera on Coh1 clones. (Coh1, wild type Coh1; Δ104, Coh1 knocked out for GBS 104, Δ647, Coh1 knocked out for sag0647, Δ648, Coh1 knocked for sag0648, Δ647-8, Coh1 knocked out for sag0647 and sag0648; 515, wild

type bacterial strain 515, which lacks an AI-1; p80 a high copy number plasmid which expresses GBS 80.) The data show that only the double sortase mutant is unable to polymerize GBS 80 indicating that the '*conditio sine qua non*' for pilus polymerization is the co-existence of GBS 80 with at least one sortase. This result leads to a reasonable assumption that SAG1405 and SAG1406 are responsible for polymerization in this strain.

Example 13: GBS 80 can be expressed in *L. lactis* under its own promoter and terminator sequences

This example demonstrates that *L. lactis*, a non-pathogenic bacterium, can express GBS AI polypeptides such as GBS 80. *L. lactis* M1363 (*J. Bacteriol. 154* (1983):1-9) was transformed with a construct encoding GBS 80. Briefly, the construct was prepared by cloning a DNA fragment containing the gene coding for GBS 80 under its own promoter and terminator sequences into plasmid pAM401 (a shuttle vector for *E. coli* and other Gram positive bacteria; *J. Bacteriol. 163* (1986):831-836). Total extracts of the transformed bacteria in log phase were separated on SDS-PAGE, transferred to membranes, and incubated with antiserum against GBS 80. A polypeptide corresponding to the molecular weight of GBS 80 was detected in the lanes containing total extracts of *L. lactis* transformed with the GBS 80 construct. See Figures 133A and 133B, lanes 6 and 7. This same polypeptide was not detected in the lane containing total extracts of *L. lactis* not transformed with the GBS 80 construct, lane 9. This example shows that *L. lactis* can express GBS 80 under its own promoter and terminator.

Example 14: *L. lactis* modified to express GBS AI-1 under the GBS 80 promoter and terminator sequences expresses GBS 80 in polymeric structures

This example demonstrates the ability of *L. lactis* to express GBS AI-1 polypeptides and to incorporate at least some of the polypeptides into oligomers. *L. lactis* was transformed with a construct containing the genes encoding GBS AI-1 polypeptides. Briefly, the construct was prepared by cloning a DNA fragment containing the genes for GBS 80, GBS 52, SAG0647, SAG0648, and GBS 104 under the GBS 80 promoter and terminator sequences into construct pAM401. The construct was transformed into *L. lactis* M1363. Total extracts of log phase transformed bacteria were separated on reducing SDS-PAGE, transferred to membranes, and incubated with antiserum against GBS 80. A polypeptide with a molecular weight corresponding to the molecular weight of GBS 80 was detected in the lanes containing *L. lactis* transformed with the GBS AI-1 encoding construct. See Figure 134, lane 2. In addition, the same lane also showed immunoreactivity of polypeptides having higher molecular weights than the polypeptide having the molecular weight of GBS 80. These higher molecular weight polypeptides are likely oligomers of GBS 80. Oligomers of similar molecular

weights were also observed on a Western blot of the culture supernatant of the transformed *L. lactis*. See lane 4 of Figure 135. Thus, this example shows that *L. lactis* transformed to express GBS AI-1 can efficiently polymerize GBS 80 in the form of a pilus. This pilus structure can likely be purified from either the cell culture supernatant or cell extracts.

Example 15: Cloning and Expression of *S. pneumoniae* Sp0462

This example describes the production of a clone encoding a Sp0462 polypeptide and expression of the clone. To produce a clone encoding Sp0462, the open reading frame encoding Sp0462 was amplified using primers that annealed within the full-length Sp0462 open reading frame sequence. Figure 150A provides a 893 amino acid sequence of Sp0462. The primers used to produce a clone encoding the Sp0462 polypeptide are shown in Figure 150B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 150A. Amplification of the open reading frame encoding Sp0462 using these primers produced the amplicon shown at lane 2 of the agarose gel provided in Figure 160. The Sp0462 clone encodes amino acid residues 38-862 of the 893 amino acid residue Sp0462 protein; the italicized residues in Figure 150A were eliminated. Figure 151A provides a schematic depiction of the recombinant Sp0462 polypeptide. Figure 151B shows a schematic depiction of the full-length Sp0462 polypeptide. Both the recombinant Sp0462 encoded by the clone and the full-length Sp0462 protein have two collagen binding protein type B (Cna B) domains and a von Hillebrand factor A (vWA) domain. The cloned recombinant Sp0462 lacks the LPXTG motif present in the full-length Sp0462 protein. Western blot analysis for expression of the Sp0462 clone did not result in detection of polypeptides with serum obtained from *S. pneumoniae*–infected patients (Figure 152A) or GBS 80 antiserum (Figure 152B).

Example 16: Cloning and Expression of *S. pneumoniae* Sp0463

This example describes the production of a clone encoding a Sp0463 polypeptide and detection of recombinant Sp0463 polypeptide expressed from the clone. To produce a clone encoding Sp0463, the open reading frame encoding Sp0463 was amplified using primers that annealed within the full-length Sp0463 open reading frame sequence. Figure 153A provides a 665 amino acid sequence of Sp0463. The primers used to produce the clone encoding Sp0463 polypeptide are shown in Figure 153B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 153A. Amplification of the open reading frame encoding Sp0463 using these primers produced the amplicon shown at lane 3 of the agarose gel provided in Figure 160. The Sp0463 clone encodes amino acid residues 23-627 of the 665 amino acid residue Sp0463 protein; the italicized residues in Figure 153A were eliminated. Figure 154A provides a schematic depiction of the recombinant Sp0463 polypeptide. Figure 154B shows a schematic depiction of the full-length Sp0463 polypeptide. Both the recombinant Sp0463 encoded by the clone and the full-length Sp0463 protein have a Cna B domain and an E box motif. The cloned recombinant

Sp0463 lacks the LPXTG motif present in the full-length Sp0463 protein. Expression of the Sp0463 clone resulted in the detection of a 60 kD polypeptide, the expected molecular weight of the recombinant Sp0463 polypeptide, by Western blot analysis. See Figure 155.

Example 17: Cloning and Expression of *S. pneumoniae* Sp0464

This example describes the production of a clone encoding a Sp0464 polypeptide and detection of recombinant Sp0464 polypeptide expressed from the clone. To produce a clone encoding Sp0464, the open reading frame encoding Sp0464 was amplified using primers that annealed either within the full-length Sp0464 open reading frame sequence. Figure 157A provides a 393 amino acid sequence of Sp0464. The primers used to produce a clone encoding the Sp0464 polypeptide are shown in Figure 157B. These primers annealed to the nucleotide sequences encoding the amino acid residues indicated by underlining in Figure 157A. Amplification of the open reading frame encoding Sp0464 using these primers produced the amplicon shown at lane 4 of the agarose gel provided in Figure 160. The Sp0464 clone encodes amino acid residues 19-356 of the 393 amino acid residue Sp0464 protein; the italicized residues in Figure 157A were eliminated. Figure 158A provides a schematic depiction of the recombinant Sp0464 polypeptide. Figure 158B shows a schematic depiction of the full-length Sp0464 polypeptide. Both the recombinant Sp0464 encoded by the clone and the full-length Sp0464 protein have two Cna B domains. The cloned recombinant Sp0464 lacks the LPXTG motif present in the full-length Sp0464 protein. Expression of the Sp0464 clone resulted in the detection of a 38 kD polypeptide, the expected molecular weight of the recombinant Sp0464 polypeptide, by Western blot analysis. See Figure 159.

Example 18: Intranasal Immunization of Mice with Recombinant *L. lactis* Expressing GBS 80 and Subsequent Challenge

This example describes a method of intranasally immunizing mice using *L. lactis* that express GBS 80. Intranasal immunization consisted of 3 doses at days 0, 14 and 28, each dose administered in three consecutive days. Each day, groups of 3 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were immunized intranasally with 10^9 or 10^{10} CFU of the recombinant *Lactococcus lactis* suspended in 20 μl of PBS. In each immunization scheme negative (wild-type *L. lactis*) and positive (recombinant GBS80) control groups were used. The immune response of the dams was monitored by using serum samples taken on day 0 and 49. The female mice were bred 2-7 days after the last immunization (at approximately t= 36 – 37), and typically had a gestation period of 21 days. Within 48 hours of birth, the pups were challenged via I.P. with GBS in a dose approximately equal to an amount which would be sufficient to kill 90 % of immunized pups (as determined by empirical data gathered from PBS control groups). The GBS challenge dose is preferably administered in 50ml of THB medium. Preferably, the pup challenge takes place at 56 to 61 days after the first immunization. The challenge inocula were prepared starting from frozen

cultures diluted to the appropriate concentration with THB prior to use. Survival of pups was monitored for 5 days after challenge.

Example 19: Subcutaneous Immunization of Mice with Recombinant *L. lactis* Expressing GBS 80 and Subsequent Challenge

This example describes a method of subcutaneous immunization mice using *L. lactis* that express GBS 80. Subcutaneous immunization consists of 3 doses at days 0, 14 and 28. Groups of 3 CD-1 outbred female mice 6-7 weeks old (Charles River Laboratories, Calco Italy) were injected subcutaneously with 10^9 or 10^{10} CFU of the recombinant *Lactococcus lactis* suspended in 100 µl of PBS. In each immunization scheme, negative (wild-type *L. lactis*) and positive (recombinant GBS80) control groups were used. The immune response of the dams was monitored by using serum samples taken on day 0 and 49. The female mice were bred 2-7 days after the last immunization (at approximately t= 36 – 37), and typically had a gestation period of 21 days. Within 48 hours of birth, the pups were challenged via I.P. with GBS in a dose approximately equal to an amount which would be sufficient to kill 90 % of immunized pups (as determined by empirical data gathered from PBS control groups). The GBS challenge dose is preferably administered in 50ml of THB medium. Preferably, the pup challenge takes place at 56 to 61 days after the first immunization. The challenge inocula were prepared starting from frozen cultures diluted to the appropriate concentration with THB prior to use. Survival of pups was monitored for 5 days after challenge.

Example 20: Immunization of Mice with GAS AI polypeptides and Subsequent Intranasal Challenge

This example describes a method of immunizing mice with GAS AI polypeptides and subsequently intranasally challenging the mice with GAS bacteria. Groups of 10 CD1 female mice aged between 6 and 7 weeks are immunized with a combination of GAS antigens of the invention GAS 15, GAS 16, and GAS 18, (15 µg of each recombinant antigen, derived from M1 strain SF370) or *L. lactis* expressing the M1 strain SF370 adhesin island, suspended in 100 µl of suitable solution. Each group receives 3 doses at days 0, 21 and 45. Immunization is performed through subcutaneous or intraperitoneal injection for the GAS 15, GAS 16, GAS 18 protein combination. The protein combination is administered with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. Immunization is performed intranasally for the *L. lactis* expressing the M1 strain SF370 adhesin island. In each immunization scheme negative and positive control groups are used.

The negative control group for the mice immunized with the GAS 15, GAS 16, GAS 18 protein combination included mice immunized with PBS. The negative control group for the mice immunized with *L. lactis* expressing the M1 strain SF370 adhesin island, included mice immunized

with either wildtype *L. lactis* or *L. lactis* transformed with the pAM401 expression vector lacking any cloned adhesin island sequence.

The positive control groups included mice immunized with purified M1 strain SF370 M protein.

Immunized mice are then anaesthetized with Zoletil and challenged intranasally with a 25 μL suspension containing 1.2 x 10^6 or 1.2 x 10^8 CFU of ISS 3348 in THB. Animals are observed daily and checked for survival.

Example 21: Active Maternal Immunization Assay

As used herein, an Active Maternal Immunization assay refers to an *in vivo* protection assay where female mice are immunized with the test antigen composition. The female mice are then bred and their pups are challenged with a lethal dose of GBS. Serum titers of the female mice during the immunization schedule are measured as well as the survival time of the pups after challenge.

Mouse immunization

Specifically, groups of 4 CD-1 outbred female mice 6-8 weeks old (Charles River Laboratories, Calco Italy) are immunized with one or more GBS antigens, (20 μg of each recombinant GBS antigen), suspended in 100 μl of PBS. Each group receives 3 doses at days 0, 21 and 35. Immunization is performed through intra-peritoneal injection of the protein with an equal volume of Complete Freund's Adjuvant (CFA) for the first dose and Incomplete Freund's Adjuvant (IFA) for the following two doses. In each immunization scheme negative and positive control groups are used.

Immune response is monitored by using serum samples taken on day 0 and 49. The sera are analyzed as pools from each group of mice.

Active maternal immunization

A maternal immunization/neonatal pup challenge model of GBS infection was used to verify the protective efficacy of the antigens in mice. The mouse protection study was adapted from Rodewald et al. (Rodewald et al. J. Infect. Diseases 166, 635 (1992)). In brief, CD-1 female mice (6-8 weeks old) were immunized before breeding, as described above. The mice received 20 μg of protein per dose when immunized with a single antigen and 60 μg of protein per dose (15 μg of each antigen) when immunized with the combination of antigens. Mice were bred 2-7 days after the last immunization. Within 48 h of birth, pups were injected intraperitoneally with 50 μl of GBS culture. Challenge inocula were prepared starting from frozen cultures diluted to the appropriate concentration with THB before use. In preliminary experiments (not shown), the challenge doses per pup for each strain tested were determined to cause 90% lethality. Survival of pups was monitored for 2 days after challenge. Protection was calculated as (percentage

deadControl minus percentage deadVaccine) divided by percentage deadControl multiplied by 100. Data were evaluated for statistical significance by Fisher's exact test.

Embodiments of the Invention

The invention encompasses, but is not limited to, the embodiments enumerated below.

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of embodiment 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

4. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide comprises a sortase substrate motif.

5. The immunogenic composition of embodiment 4 wherein the sortase substrate motif is an LPXTG motif.

6. The immunogenic composition of embodiment 5 wherein the LPXTG motif is represented by the amino acid sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

7. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

8. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

9. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

10. The immunogenic composition of any one of embodiments 1-3 wherein the GBS AI polypeptide is capable of associating with an epithelial cell surface.

11. The immunogenic composition of embodiment 10 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

12. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide is a full-length GBS AI protein.

13. The immunogenic composition of any of embodiments 1-3 wherein the GBS AI polypeptide is a fragment of a full-length GBS AI protein.

14. The immunogenic composition of embodiment 13 wherein the fragment comprises at least 7 contiguous amino acid residues of the GBS AI protein.

15. The immunogenic composition of embodiment 2 wherein the GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

16. The immunogenic composition of embodiment 3 wherein the GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

17. The immunogenic composition of embodiment 15 wherein the GBS AI polypeptide is GBS 80.

18. The immunogenic composition of any of embodiments 1-3 or 15-17 wherein the oligomeric form is a hyperoligomer.

19. The immunogenic composition of any of embodiments 1-3, or 15-17 further comprising a Gram positive bacterium antigen not associated with an AI.

20. The immunogenic composition of embodiment 19 wherein the antigen is selected from the group consisting of GBS 322 and GBS 276.

21. The immunogenic composition of embodiment 20 wherein the antigen is GBS 322.

22. An immunogenic composition comprising a purified Gram positive bacteria adhesin island (AI) polypeptide in an oligomeric form.

23. The immunogenic composition of embodiment 22 wherein the Gram positive bacteria is of a genus selected from the group consisting of *Streptococcus*, *Enterococcus*, *Staphylococcus*, or *Listeria*.

24. The immunogenic composition of embodiment 23 wherein the Gram positive bacteria is of the genus *Streptococcus*.

25. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide comprises a sortase substrate motif.

26. The immunogenic composition of embodiment 25 wherein the sortase substrate motif is an LPXTG motif.

27. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to adhere to epithelial cells.

28. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to invade epithelial cells.

29. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide affects the ability of Gram positive bacteria to translocate through an epithelial cell layer.

30. The immunogenic composition of any one of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is capable of associating with an epithelial cell surface.

31. The immunogenic composition of embodiment 30 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

32. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is a full-length Gram positive bacteria AI protein.

33. The immunogenic composition of any of embodiments 22-24 wherein the Gram positive bacteria AI polypeptide is a fragment of a full-length Gram positive bacteria AI protein.

34. The immunogenic composition of embodiment 33 wherein the fragment comprises at least 7 contiguous amino acid residues of the Gram positive bacteria AI protein.

35. The immunogenic composition of embodiment 24 wherein the genus *Streptococcus* bacteria is Group A Streptococcus (GAS) bacteria and the Gram positive bacteria AI polypeptide is a GAS AI polypeptide.

36. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-1.

37. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-2.

38. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-3.

39. The immunogenic composition of embodiment 35 wherein the GAS AI polypeptide is selected from a GAS AI-4.

40. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide comprises a sortase substrate motif.

41. The immunogenic composition of embodiment 40 wherein the sortase substrate motif is an LPXTG motif.

42. The immunogenic composition of embodiment 41 wherein the LPXTG motif is represented by XXXXG, wherein the X at the first amino acid position is an L, a V, an E, or a Q, wherein the X at the second amino acid position is P if the X at the first amino acid position is an L, the X at the second amino acid position is a V if the X at the first amino acid position is an E or a Q, or the X at the second amino acid position is a V or a P if the X at the first amino acid position is a V, wherein the X at the third amino acid position is any amino acid residue, and wherein the X at the fourth amino acid position is a T if the X at the first amino acid position is a V, an E, or a Q, or the X at the fourth amino acid position is a T, an S, or an A if the X at the first amino acid position is an L.

43. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to adhere to epithelial cells.

44. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to invade epithelial cells.

45. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide affects the ability of GAS bacteria to translocate through an epithelial cell layer.

46. The immunogenic composition of any one of embodiments 35-39 wherein the GAS AI polypeptide is capable of associating with an epithelial cell surface.

47. The immunogenic composition of embodiment 46 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

48. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide is a full-length GAS AI protein.

49. The immunogenic composition of any of embodiments 35-39 wherein the GAS AI polypeptide is a fragment of a full-length GAS AI protein.

50. The immunogenic composition of embodiment 49 wherein the fragment comprises at least 7 contiguous amino acid residues of the GAS AI protein.

51. The immunogenic composition of embodiment 36 wherein the GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

52. The immunogenic composition of embodiment 37 wherein the GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

53. The immunogenic composition of embodiment 38 wherein the GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

53. The immunogenic composition of embodiment 39 wherein the GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

54. The immunogenic composition of embodiment 24 wherein the *Streptococcus* bacteria is *Streptococcus pneumoniae* and the Gram positive bacteria AI polypeptide is a *S. pneumoniae* AI polypeptide.

55. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide comprises a sortase substrate motif.

56. The immunogenic composition of embodiment 55 wherein the sortase substrate motif is an LPXTG motif.

57. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to adhere to epithelial cells.

58. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to invade epithelial cells.

59. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to translocate through an epithelial cell layer.

60. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is capable of associating with an epithelial cell surface.

61. The immunogenic composition of embodiment 60 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

62. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is a full-length *S. pneumoniae* AI protein.

63. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is a fragment of a full-length *S. pneumoniae* AI protein.

64. The immunogenic composition of embodiment 63 wherein the fragment comprises at least 7 contiguous amino acid residues of the *S. pneumoniae* AI protein.

65. The immunogenic composition of embodiment 54 wherein the *S. pneumoniae* AI polypeptide is selected from the group consisting of SP0462, SP0463, SP0464, orf3_670, orf4_670, orf5_670, ORF3_14CSR, ORF4_14CSR, ORF5_14CSR, ORF3_19AH, ORF4_19AH, ORF5_19AH, ORF3_19FTW, ORF4_19FTW, ORF5_19FTW, ORF3_23FP, ORF4_23FP, ORF5_23FP, ORF3_23FTW, ORF4_23FTW, ORF5_23FTW, ORF3_6BF, ORF4_6BF, ORF5_6BF, ORF3_6BSP, ORF4_6BSP, ORF5_6BSP, ORF3_9VSP, ORF4_9VSP, ORF5_9VSP, and fragments thereof.

66. The immunogenic composition of any one of embodiments 22-24, 35-39, 51-54, or 65 wherein the oligomeric form is a hyperoligomer.

67. The immunogenic composition of any one of embodiments 22-24, 35-39, 51-54, or 65 further comprising a Gram positive bacteria antigen not associated with an AI.

68. The immunogenic composition of embodiment 67 wherein the antigen is selected from the group consisting of GBS 322 and GBS 276.

69. An immunogenic composition comprising a first and a second Group B Streptococcus (GBS) adhesin island (AI) polypeptide.

70. The immunogenic composition of embodiment 69 wherein a full-length polynucleotide sequence encoding for the first GBS AI polypeptide is not present in a GBS bacteria genome comprising a polynucleotide sequence encoding for the second GBS AI polypeptide.

71. The immunogenic composition of embodiment 69 wherein polynucleotides encoding the first and the second GBS AI polypeptide are each present in genomes of more than one GBS serotype and strain isolate.

72. The immunogenic composition of embodiment 69 wherein the first GBS AI polypeptide is encoded by a GBS AI-1.

73. The immunogenic composition of embodiment 69 wherein the first GBS AI polypeptide is encoded by a GBS AI-2.

74. The immunogenic composition of embodiment 72 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

75. The immunogenic composition of embodiment 73 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

76. The immunogenic composition of embodiment 72 wherein the second GBS AI polypeptide is encoded by a GBS AI-1.

77. The immunogenic composition of embodiment 73 wherein the second GBS AI polypeptide is encoded by a GBS AI-1.

78. The immunogenic composition of embodiment 72 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

79. The immunogenic composition of embodiment 73 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

80. The immunogenic composition of embodiment 74 or 75 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

81. The immunogenic composition of embodiment 76 or 77 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

82. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide comprises a sortase substrate motif.

83. The immunogenic composition of embodiment 82 wherein the sortase substrate motif is an LPXTG motif.

84. The immunogenic composition of embodiment 83 wherein the LPXTG motif is represented by the sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

85. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

86. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

87. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

88. The immunogenic composition of any one of embodiments 69-77 wherein the first GBS AI polypeptide is capable of associating with an epithelial cell surface.

89. The immunogenic composition of embodiment 88 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

90. The immunogenic composition of any of embodiments 69-77 wherein the first GBS AI polypeptide is a full-length GBS AI protein.

91. The immunogenic composition of any of embodiments 69-77 wherein the first GBS AI polypeptide is a fragment of a full-length GBS AI protein.

92. The immunogenic composition of embodiment 91 wherein the fragment comprises at least 7 contiguous amino acid residues of the first GBS AI protein.

93. The immunogenic composition of any one of embodiments 69-79 wherein the first GBS AI polypeptide is in oligomeric form.

94. The immunogenic composition of any one of embodiments 69-77 wherein the second GBS AI polypeptide is in oligomeric form.

95. The immunogenic composition of any one of embodiments 69-79 wherein the first and the second GBS AI polypeptide are associated in a single oligomeric form.

96. The immunogenic composition of embodiment 95 wherein the first and the second GBS AI polypeptides are chemically associated.

97. The immunogenic composition of embodiment 95 wherein the first and the second GBS AI polypeptides are physically associated.

98. The immunogenic composition of embodiment 93 wherein the oligomeric form is a hyperoligomer.

99. The immunogenic composition of embodiment 94 wherein the oligomeric form is a hyperoligomer.

100. The immunogenic composition of embodiment 76 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 104.

101. The immunogenic composition of embodiment 74 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 67.

102. The immunogenic composition of any one of embodiments 69-79, 100, or 101 further comprising a GBS polypeptide not associated with an AI.

103. The immunogenic composition of embodiment 102 wherein the GBS polypeptide not associated with an AI is selected from the group consisting of GBS 322 and GBS 276.

104. The immunogenic composition of embodiment 103 wherein the GBS polypeptide not associated with an AI is GBS 322.

105. An immunogenic composition comprising a first and a second Gram positive bacteria adhesin island (AI) polypeptide.

106. The immunogenic composition of embodiment 105 wherein a full length polynucleotide sequence encoding for the first Gram positive bacteria AI polypeptide is not present in a genome of a Gram positive bacteria comprising a full length polynucleotide sequence encoding for the second Gram positive bacteria AI polypeptide.

107. The immunogenic composition of embodiment 105 wherein polynucleotides encoding the first and the second Gram positive bacteria AI polypeptide are each present in genomes of more than one Gram positive bacteria serotype and strain isolate.

108. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are of different Gram positive bacteria species.

109. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are of the same Gram positive bacteria species.

110. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are from different AI subtypes.

111. The immunogenic composition of embodiment 105 wherein the first and the second Gram positive bacteria AI polypeptides are from the same AI subtype.

112. The immunogenic composition of embodiment 105 wherein the first Gram positive bacteria AI polypeptide has detectable surface exposure on a first Gram positive bacteria strain or serotype but not a second Gram positive bacteria strain or subtype and the second Gram positive bacteria AI polypeptide has detectable surface exposure on the second Gram positive bacteria strain or serotype but not the first Gram positive bacteria strain or serotype.

113. The immunogenic composition of embodiment 105 wherein the Gram positive bacteria is *S. pneumonaie*, *S. mutans*, *E. faecalis*, *E. faecium*, *C. difficile*, *L. monocytogenes*, or *C. diphtheriae*.

114. The immunogenic composition of any of embodiments 105-113 wherein the first and the second Gram positive bacteria AI polypeptides comprise a sortase substrate motif.

115. The immunogenic composition of embodiment 114 wherein the sortase substrate motif is an LPXTG motif.

116. The immunogenic composition of embodiment 115 wherein the LPXTG motif is represented by XXXXG, wherein the X at amino acid position 1 is an L, a V, an E, an I, an F, or a Q, wherein X at amino acid position 2 is a P if X at amino acid position 1 is an L, an I, or an F, wherein X at amino acid position 2 is a V if X at amino acid position 1 is a E or a Q, wherein X at amino acid position 2 is a V or a P if X at amino acid position 1 is a V, wherein X at amino acid position 3 is any amino acid residue, wherein X at amino acid position 4 is a T if X at amino acid position 1 is a V, E, I, F, or Q, and wherein X at amino acid position 4 is a T, S, or A if X at amino acid position 1 is an L.

117. The immunogenic composition of embodiment 105 wherein the first Gram positive bacteria AI polypeptide is a first Group A Streptococcus (GAS) AI polypeptide.

118. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide comprises a sortase substrate motif.

119. The immunogenic composition of embodiment 118 wherein the sortase substrate motif is an LPXTG motif.

120. The immunogenic composition of embodiment 119 wherein the LPXTG motif is represented by XXXXG, wherein the X at the first amino acid position is an L, a V, an E, or a Q, wherein the X at the second amino acid position is P if the X at the first amino acid position is an L, the X at the second amino acid position is a V if the X at the first amino acid position is an E or a Q, or the X at the second amino acid position is a V or a P if the X at the first amino acid position is a V, wherein the X at the third amino acid position is any amino acid residue, and wherein the X at the fourth amino acid position is a T if the X at the first amino acid position is a V, an E, or a Q, or the X at the fourth amino acid position is a T, an S, or an A if the X at the first amino acid position is an L.

121. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to adhere to epithelial cells.

122. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to invade epithelial cells.

123. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide affects the ability of GAS bacteria to translocate through an epithelial cell layer.

124. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is capable of associating with an epithelial cell surface.

125. The immunogenic composition of embodiment 117 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

126. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a full-length GAS AI protein.

127. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a fragment of a full-length GAS AI protein.

128. The immunogenic composition of embodiment 127 wherein the fragment comprises at least 7 contiguous amino acid residues of the GAS AI protein.

129. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-1 polypeptide.

130. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-2 polypeptide.

131. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-3 polypeptide.

132. The immunogenic composition of embodiment 117 wherein the first GAS AI polypeptide is a first GAS AI-4 polypeptide.

133. The immunogenic composition of any one of embodiments 117 or 129-132 wherein the second Gram positive bacteria AI polypeptide is a second GAS AI polypeptide.

134. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-1 polypeptide.

135. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-2 polypeptide.

136. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-3 polypeptide.

137. The immunogenic composition of embodiment 133 wherein the second GAS AI polypeptide is a second GAS AI-4 polypeptide.

138. The immunogenic composition of embodiment 129 wherein the first GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

139. The immunogenic composition of embodiment 130 wherein the first GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

140. The immunogenic composition of embodiment 131 wherein the first GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

141. The immunogenic composition of embodiment 132 wherein the first GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

142. The immunogenic composition of embodiment 134 wherein the second GAS AI-1 polypeptide is selected from the group consisting of M6_Spy0157, M6_Spy0159, M6_Spy0160, CDC SS 410_fimbrial, ISS3650_fimbrial, DSM2071_fimbrial, and fragments thereof.

143. The immunogenic composition of embodiment 135 wherein the second GAS AI-2 polypeptide is selected from the group consisting of GAS15, GAS16, GAS18, and fragments thereof.

144. The immunogenic composition of embodiment 136 wherein the second GAS AI-3 polypeptide is selected from the group consisting of SpyM3_0098, SpyM3_0100, SpyM3_0102, SpyM3_0104, SPs0100, SPs0102, SPs0104, SPs0106, orf78, orf80, orf82, orf84, spyM18_0126, spyM18_0128, spyM18_0130, spyM18_0132, SpyoM01000156, SpyoM01000155, SpyoM01000154, SpyoM01000153, SpyoM01000152, SpyoM01000151, SpyoM01000150, SpyoM01000149, ISS3040_fimbrial, ISS3776_fimbrial, ISS4959_fimbrial, and fragments thereof.

145. The immunogenic composition of embodiment 137 wherein the second GAS AI-4 polypeptide is selected from the group consisting of 19224134, 19224135, 19224137, 19224139, 19224141, 20010296_fimbrial, 20020069_fimbrial, CDC SS 635_fimbrial, ISS4883_fimbrial, ISS4538_fimbrial, and fragments thereof.

146. The immunogenic composition of any one of embodiments 117-132 or 138-141 wherein the second Gram positive bacteria AI polypeptide is a Group B Streptococcus (GBS) AI polypeptide.

147. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide comprises a sortase substrate motif.

148. The immunogenic composition of embodiment 147 wherein the sortase substrate motif is an LPXTG motif.

149. The immunogenic composition of embodiment 148 wherein the LPXTG motif is represented by the amino acid sequence XPXTG, wherein the X at amino acid position 1 is an L, an I, or an F and the X at amino acid position 3 is any amino acid residue.

150. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to adhere to epithelial cells.

151. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to invade epithelial cells.

152. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide affects the ability of GBS bacteria to translocate through an epithelial cell layer.

153. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is capable of associating with an epithelial cell surface.

154. The immunogenic composition of embodiment 146 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

155. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a full-length GBS AI protein.

156. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a fragment of a full-length GBS AI protein.

157. The immunogenic composition of embodiment 156 wherein the fragment comprises at least 7 contiguous amino acid residues of the GBS AI protein.

158. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a GBS AI-1 polypeptide.

159. The immunogenic composition of embodiment 146 wherein the GBS AI polypeptide is a GBS AI-2 polypeptide.

160. The immunogenic composition of embodiment 158 wherein the GBS AI-1 polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

161. The immunogenic composition of embodiment 159 wherein the GBS AI-2 polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

162. The immunogenic composition of any one of embodiments 117-132 or 138-141 wherein the second Gram positive bacteria AI polypeptide is a *Streptococcus pneumoniae* AI polypeptide.

163. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide comprises a sortase substrate motif.

164. The immunogenic composition of embodiment 163 wherein the sortase substrate motif is an LPXTG motif.

165. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to adhere to epithelial cells.

166. The immunogenic composition of embodiment 162 *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to invade epithelial cells.

167. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide affects the ability of *S. pneumoniae* to translocate through an epithelial cell layer.

168. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is capable of associating with an epithelial cell surface.

169. The immunogenic composition of embodiment 168 wherein the associating with an epithelial cell surface is binding to the epithelial cell surface.

170. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is a full-length *S. pneumoniae* AI protein.

171. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is a fragment of a full-length *S. pneumoniae* AI protein.

172. The immunogenic composition of embodiment 162 wherein the fragment comprises at least 7 contiguous amino acid residues of the *S. pneumoniae* AI protein.

173. The immunogenic composition of embodiment 162 wherein the *S. pneumoniae* AI polypeptide is selected from the group consisting of SP0462, SP0463, SP0464, orf3_670, orf4_670, orf5_670, ORF3_14CSR, ORF4_14CSR, ORF5_14CSR, ORF3_19AH, ORF4_19AH, ORF5_19AH, ORF3_19FTW, ORF4_19FTW, ORF5_19FTW, ORF3_23FP, ORF4_23FP, ORF5_23FP, ORF3_23FTW, ORF4_23FTW, ORF5_23FTW, ORF3_6BF, ORF4_6BF, ORF5_6BF, ORF3_6BSP, ORF4_6BSP, ORF5_6BSP, ORF3_9VSP, ORF4_9VSP, ORF5_9VSP, and fragments thereof.

174. The immunogenic composition of any one of embodiments 105-117 wherein the first Gram positive bacteria AI polypeptide is in oligomeric form.

175. The immunogenic composition of embodiment 174 wherein the oligomeric form is a hyperoligomer.

176. The immunogenic composition of embodiment 174 wherein the second Gram positive bacteria AI polypeptide is in oligomeric form.

177. The immunogenic composition of embodiment 176 wherein the oligomeric form is a hyperoligomer.

178. The immunogenic composition of embodiment 176 wherein the first and the second Gram positive bacteria AI polypeptide are associated in a single oligomeric form.

179. The immunogenic composition of embodiment 178 wherein the first and the second Gram positive bacteria AI polypeptide are chemically associated.

180. The immunogenic composition of embodiment 178 wherein the first and the second Gram positive bacteria AI polypeptide are physically associated.

181. The immunogenic composition of any one of embodiments 105-117 further comprising a Gram positive bacteria polypeptide not associated with an AI.

182. The immunogenic composition of embodiment 181 wherein the Gram positive bacteria polypeptide not associated with an AI is selected from the group consisting of GBS 322 and GBS 276.

183. The immunogenic composition of embodiment 182 wherein the Gram positive bacteria polypeptide not associated with an AI is GBS 322.

184. A modified Gram positive bacterium adapted to produce increased levels of AI surface protein.

185. The modified Gram positive bacterium of embodiment 184 wherein the AI surface protein is in oligomeric form.

186. The modified Gram positive bacterium of embodiment 185 wherein the oligomeric form is a hyperoligomer.

187. The modified Gram positive bacterium of any one of embodiments 184-186 which is a Group B Streptococcus bacterium.

188. The modified Gram positive bacterium of any one of embodiments 184-186 which is a Group A Streptococcus bacterium.

189. The modified Gram positive bacterium of any one of embodiments 184-186 which is a non-pathogenic Gram positive bacterium.

190. The modified Gram positive bacterium of embodiment 189 wherein the non-pathogenic Gram positive bacterium is *Streptococus gordonii.*

191. The modified Gram positive bacterium of embodiment 189 wherein the non-pathogenic Gram positive bacterium is *Lactococcus lactis*.

192. The modified Gram positive bacterium of any one of embodiments 184-186 which has been inactivated and wherein the AI surface protein is exposed on the surface of the Gram positive bacterium.

193. The modified Gram positive bacterium of any one of embodiments 184-186 which has been attenuated and wherein the AI surface protein is exposed on the surface of the Gram positive bacterium.

194. The modified GBS bacterium of embodiment 187 which has been inactivated and wherein the AI surface protein is exposed on the surface of the GBS bacterium.

195. The modified GBS bacterium of embodiment 187 which has been attenuated and wherein the AI surface protein is exposed on the surface of the GBS bacterium.

196. The modified GAS bacterium of embodiment 188 which has been inactivated and wherein the AI surface protein is exposed on the surface of the GAS bacterium.

197. The modified GAS bacterium of embodiment 188 which has been attenuated and wherein the AI surface protein is exposed on the surface of the GAS bacterium.

198. The modified non-pathogenic bacterium of embodiment 189 which has been inactivated and wherein the AI surface protein is exposed on the surface of the non-pathogenic Gram positive bacterium.

199. The modified non-pathogenic bacterium of embodiment 189 which has been attenuated and wherein the AI surface protein is exposed on the surface of the non-pathogenic Gram positive bacterium.

200. A method for manufacturing an oligomeric adhesin island (AI) surface antigen comprising:

culturing a Gram positive bacterium that expresses an oligomeric AI surface antigen and isolating the expressed oligomeric AI surface antigen.

201. The method of embodiment 200 wherein the step of isolating is performed by collecting said oligomeric AI surface antigen from Gram positive bacterium secretions in the Gram positive bacterium culture.

202. The method of embodiment 200 further comprising a step of purifying.

203. The method of embodiment 202 wherein the oligomeric AI surface antigen is purified from the Gram positive bacterium cell surface.

204. The method of embodiment 200 wherein the Gram positive bacterium is adapted for increased AI protein expression.

205. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is a Group A Streptococcus bacterium.

206. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is a Group B Streptococcus bacterium.

207. The method of any one of embodiments 200-204 wherein the oligomeric AI surface antigen is in hyperoligomeric form.

208. The method of embodiment 200 wherein the Gram positive bacterium expresses the oligomeric AI surface antigen recombinantly.

209. The method of embodiment 208 wherein the Gram positive bacterium further manipulated expresses at least 1 AI sortase.

210. The modified Gram positive bacterium of any one of embodiments 184-186 which is a *S. pneumoniae* bacterium.

211. The method of any one of embodiments 200-204 wherein the Gram positive bacterium is *S. pneumoniae*.

WE CLAIM:

1. An immunogenic composition comprising a purified Group B Streptococcus (GBS) adhesin island (AI) polypeptide in oligomeric form.

2. The immunogenic composition of claim 1 wherein the GBS AI polypeptide is selected from a GBS AI-1.

3. The immunogenic composition of claim 1 wherein the GBS AI polypeptide is selected from a GBS AI-2.

4. The immunogenic composition of claim 2 wherein the GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

5. The immunogenic composition of claim 3 wherein the GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof.

6. The immunogenic composition of claim 4 wherein the GBS AI polypeptide is GBS 80.

7. The immunogenic composition of any of claims 1-6 wherein the oligomeric form is a hyperoligomer.

8 (22). An immunogenic composition comprising a purified Gram positive bacteria adhesin island (AI) polypeptide in an oligomeric form.

9 (23). The immunogenic composition of claim 8 wherein the Gram positive bacteria is of a genus selected from the group consisting of *Streptococcus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, *Corynebacterium*, or *Listeria*.

10 (24). The immunogenic composition of claim 9 wherein the Gram positive bacteria is of the genus *Streptococcus*.

11 (35). The immunogenic composition of claim 10 wherein the genus *Streptococcus* bacteria is Group A Streptococcus (GAS) bacteria and the Gram positive bacteria AI polypeptide is a GAS AI polypeptide.

12 (36). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-1.

13 (37). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-2.

14 (38). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-3.

15 (39). The immunogenic composition of claim 11 wherein the GAS AI polypeptide is selected from a GAS AI-4.

16 (66). The immunogenic composition of any one of claims 8-15 wherein the oligomeric form is a hyperoligomer.

17. An immunogenic composition comprising a first and a second Group B Streptococcus (GBS) adhesin island (AI) polypeptide.

18. The immunogenic composition of claim 17 wherein the first GBS AI polypeptide is encoded by a GBS AI-1.

19. The immunogenic composition of claim 18 wherein the second GBS AI polypeptide is encoded by a GBS AI-2.

20. The immunogenic composition of claim 18 wherein the first GBS AI polypeptide is selected from the group consisting of GBS 80, GBS 104, GBS 52, and fragments thereof.

21. The immunogenic composition of claim 19 wherein the second GBS AI polypeptide is selected from the group consisting of GBS 59, GBS 67, GBS 150, 01521, 01523, 01524, and fragments thereof, and wherein the first and the second GBS AI polypeptide are not the same polypeptide.

22. The immunogenic composition of claim 19 wherein the first GBS AI polypeptide is GBS 80 and the second GBS AI polypeptide is GBS 67.

23. An immunogenic composition comprising a first and a second Gram positive bacteria adhesin island (AI) polypeptide.

24. The immunogenic composition of claim 23 wherein the Gram positive bacteria is *Streptococcus*, *Enterococcus*, *Staphylococcus*, *Clostridium*, *Corynebacterium*, or *Listeria*.

25. The immunogenic composition of claim 23 wherein the first Gram positive bacteria AI polypeptide is a first Group A Streptococcus (GAS) AI polypeptide.

26. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-1 polypeptide.

27. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-2 polypeptide.

28. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-3 polypeptide.

29. The immunogenic composition of claim 25 wherein the first GAS AI polypeptide is a first GAS AI-4 polypeptide.

30. The immunogenic composition of any one of claims 25-29 wherein the second Gram positive bacteria AI polypeptide is a second GAS AI polypeptide.

31. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-1 polypeptide.

32. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-2 polypeptide.

33. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-3 polypeptide.

34. The immunogenic composition of claim 30 wherein the second GAS AI polypeptide is a second GAS AI-4 polypeptide.

35. A modified Gram positive bacterium adapted to produce increased levels of AI surface protein.

36. The modified Gram positive bacterium of claim 35 wherein the AI surface protein is in oligomeric form.

37. The modified Gram positive bacterium of claim 36 wherein the oligomeric form is a hyperoligomer.

38. The modified Gram positive bacterium of any one of claims 35-37 which is a non-pathogenic Gram positive bacterium.

39. The modified Gram positive bacterium of claim 38 wherein the non-pathogenic Gram positive bacterium is *Lactococcus lactis.*

40. A method for manufacturing an oligomeric adhesin island (AI) surface antigen comprising:

culturing a Gram positive bacterium that expresses an oligomeric AI surface antigen and isolating the expressed oligomeric AI surface antigen.

FIGURE 1: Adhesion Island 1

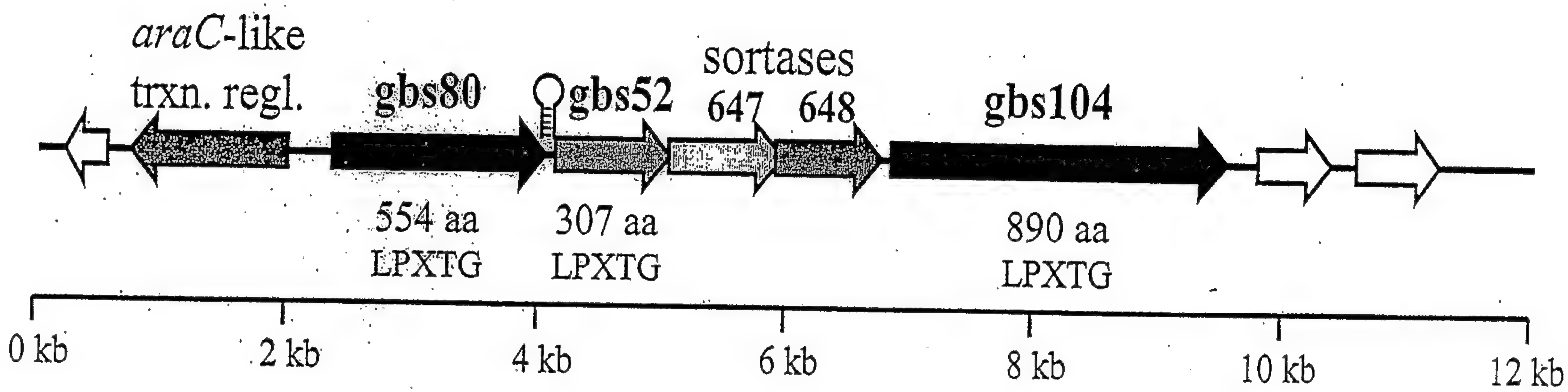


Figure 2: Conservation of AI-1 in GBS serotypes and strain isolates

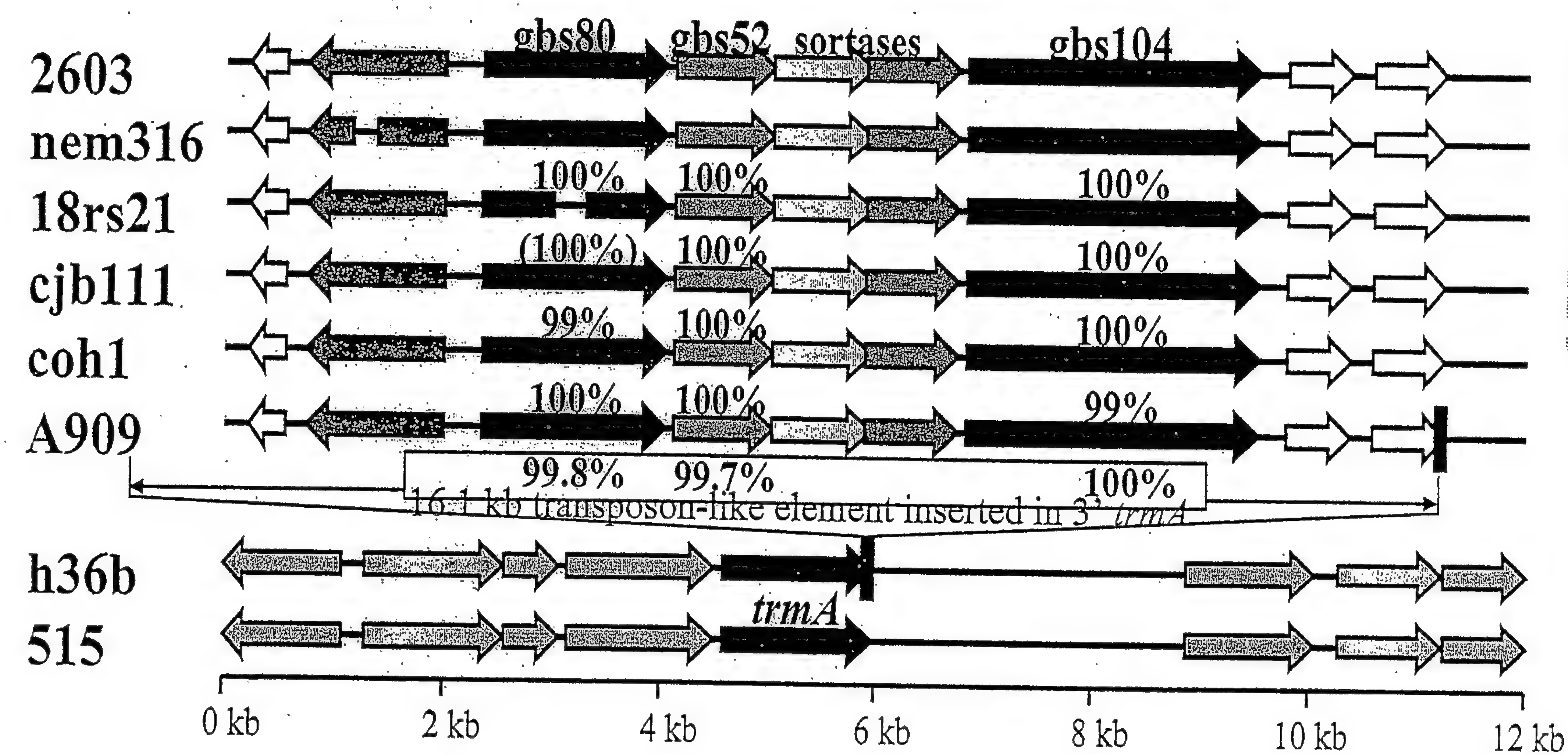


FIGURE 3: Correlation of AI-1 and AI-2 within GBS serotype V, strain isolate 2603 genome

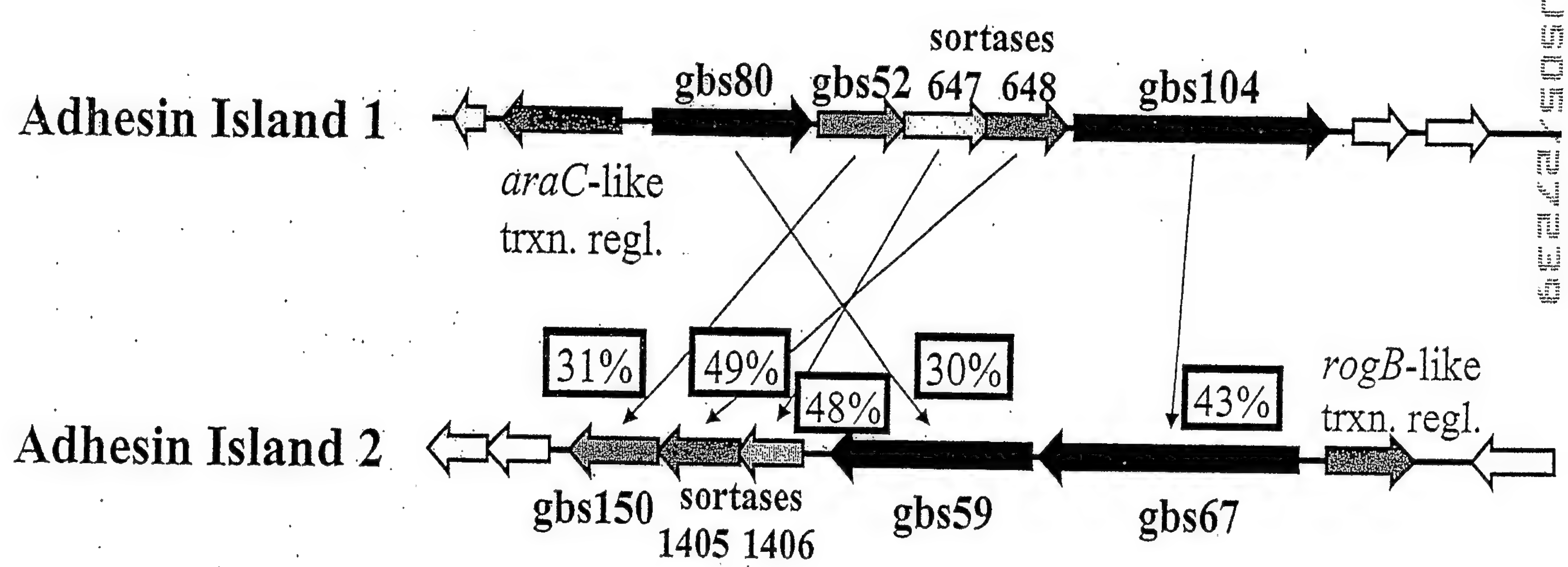


Figure 4: Identification and Variance of AI-2 in Several GBS Serotypes and Strain Isolates

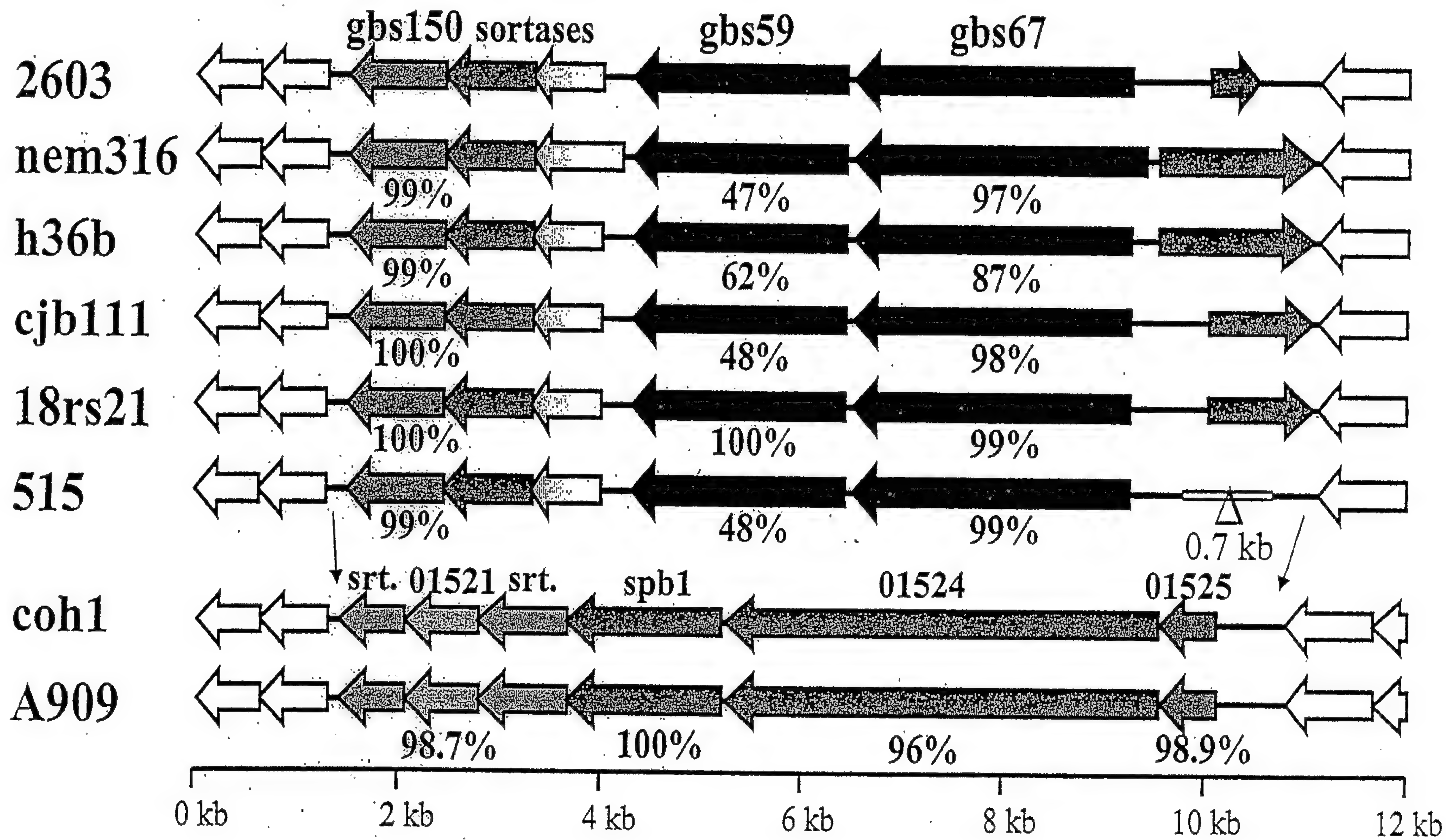
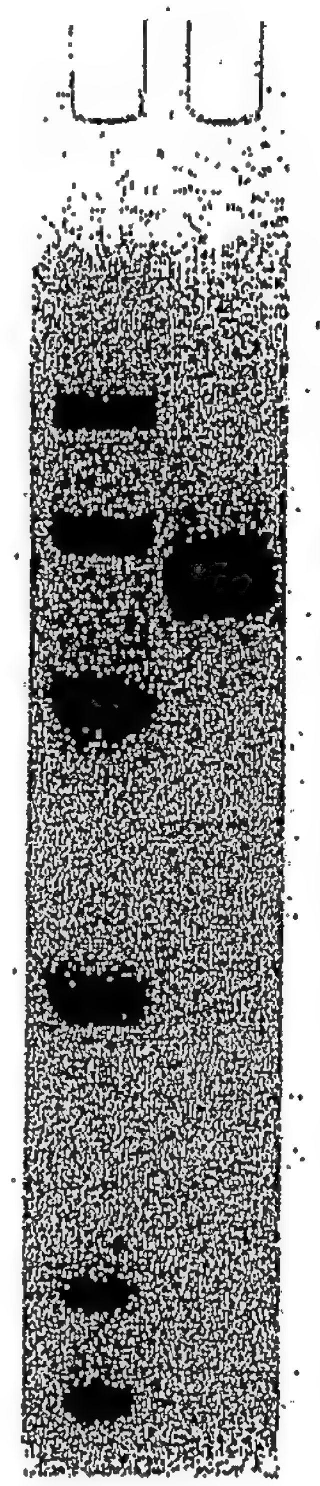


Figure 5: Purified gbs80 protein binds fibronectin and fibrinogen in an ELISA

A:



B:

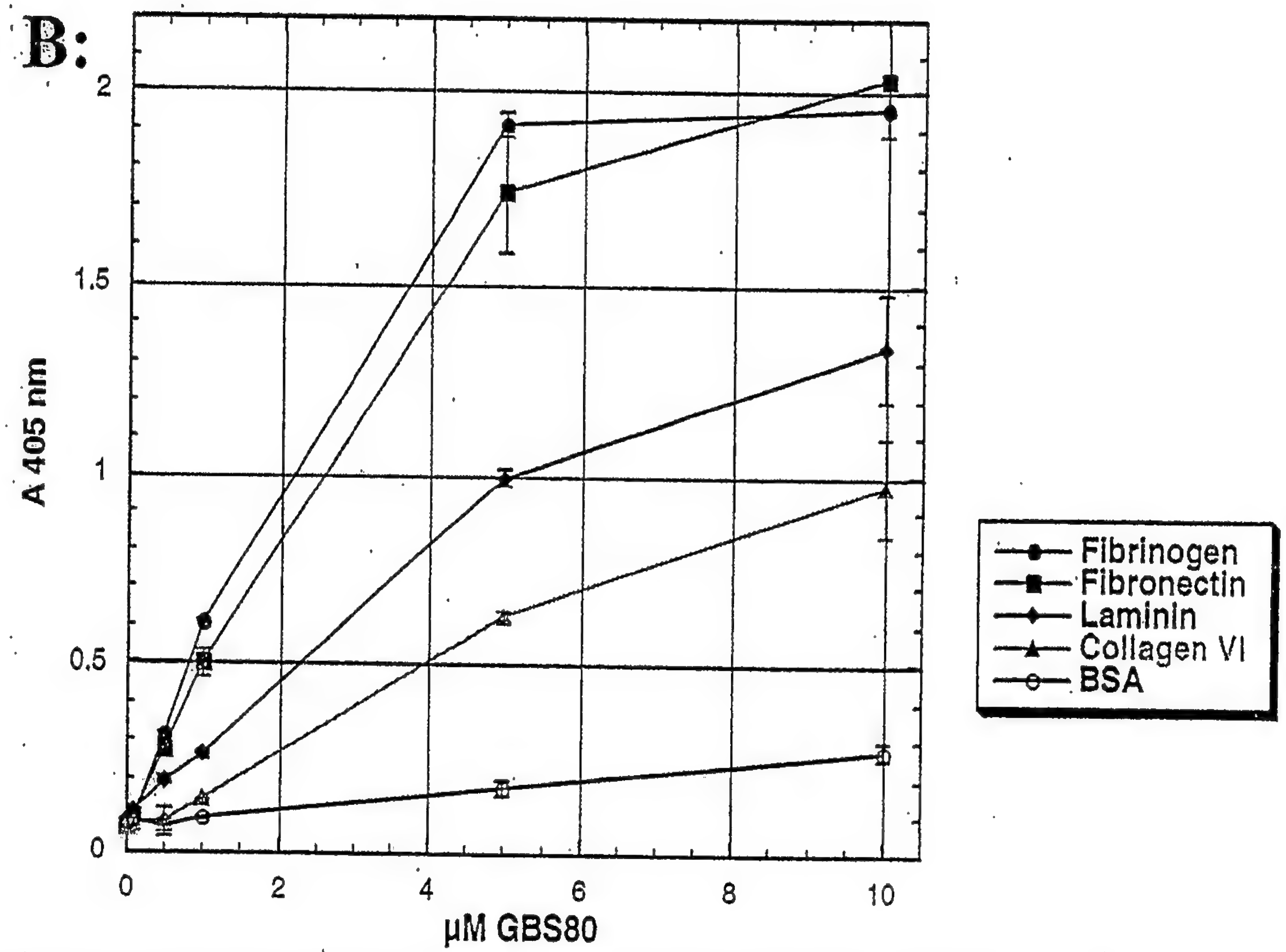


Figure 6: Adhesion Island I is an operon by RT-PCR

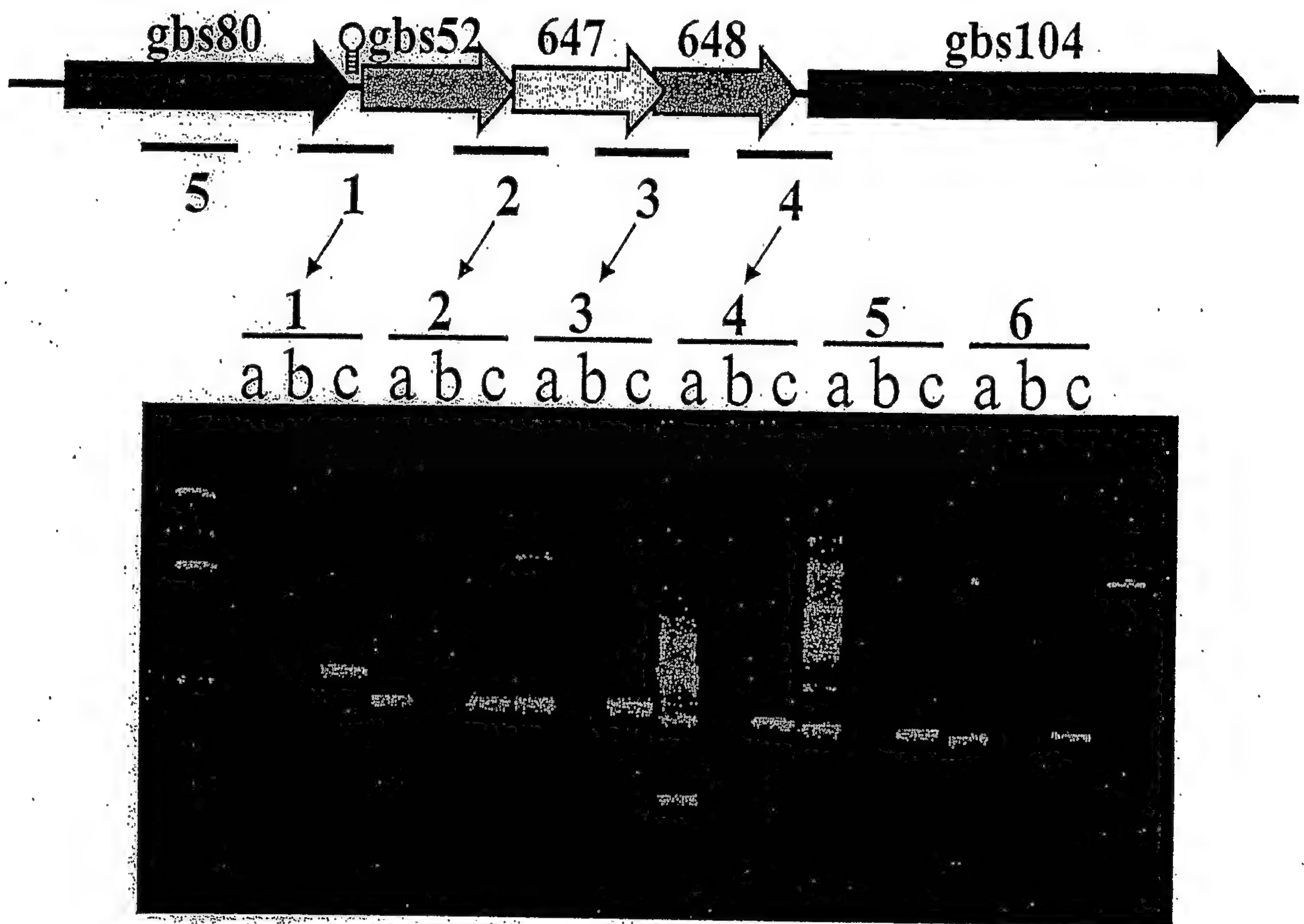
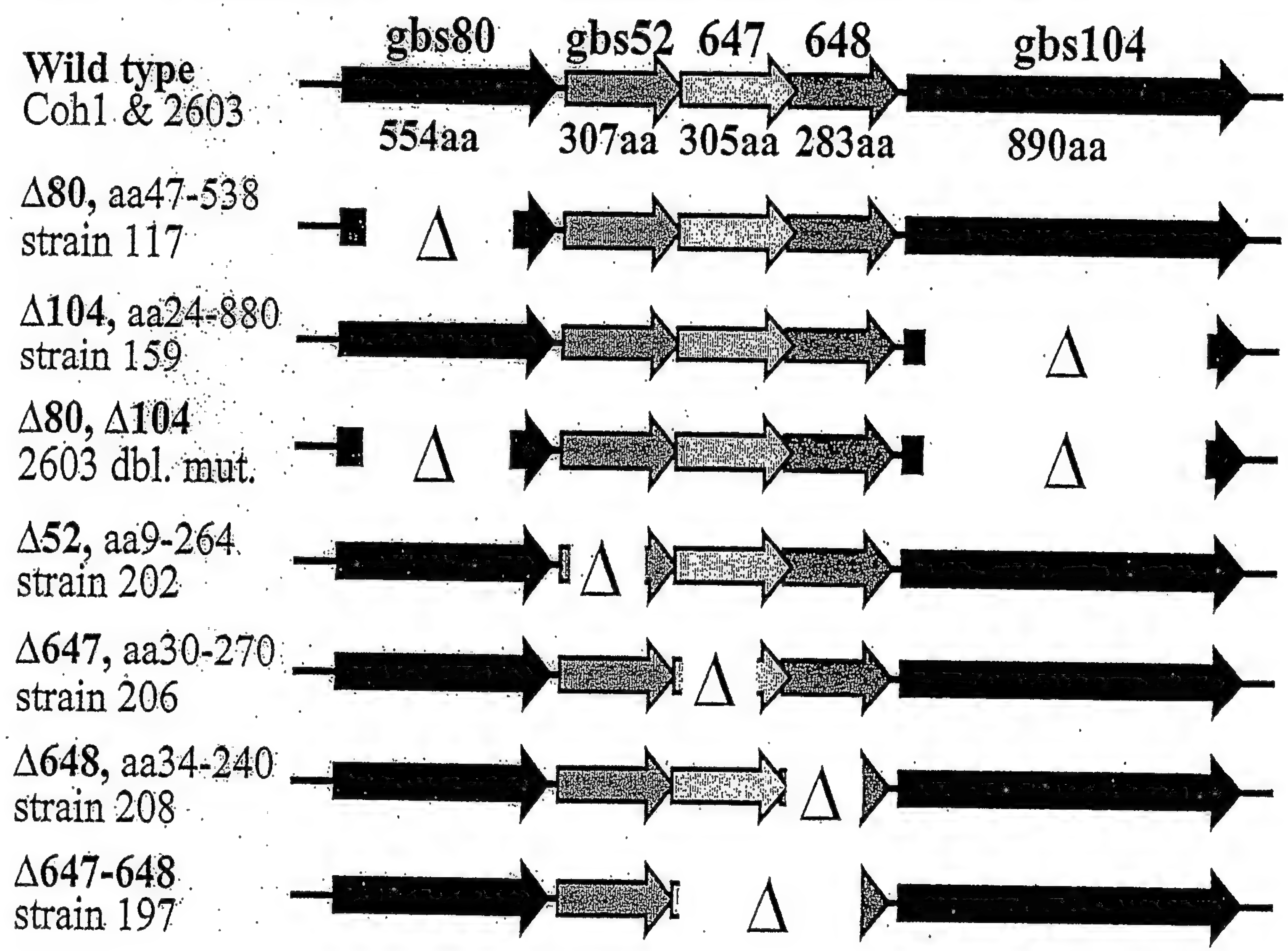


Figure 7: In frame deletions of AI-1 genes



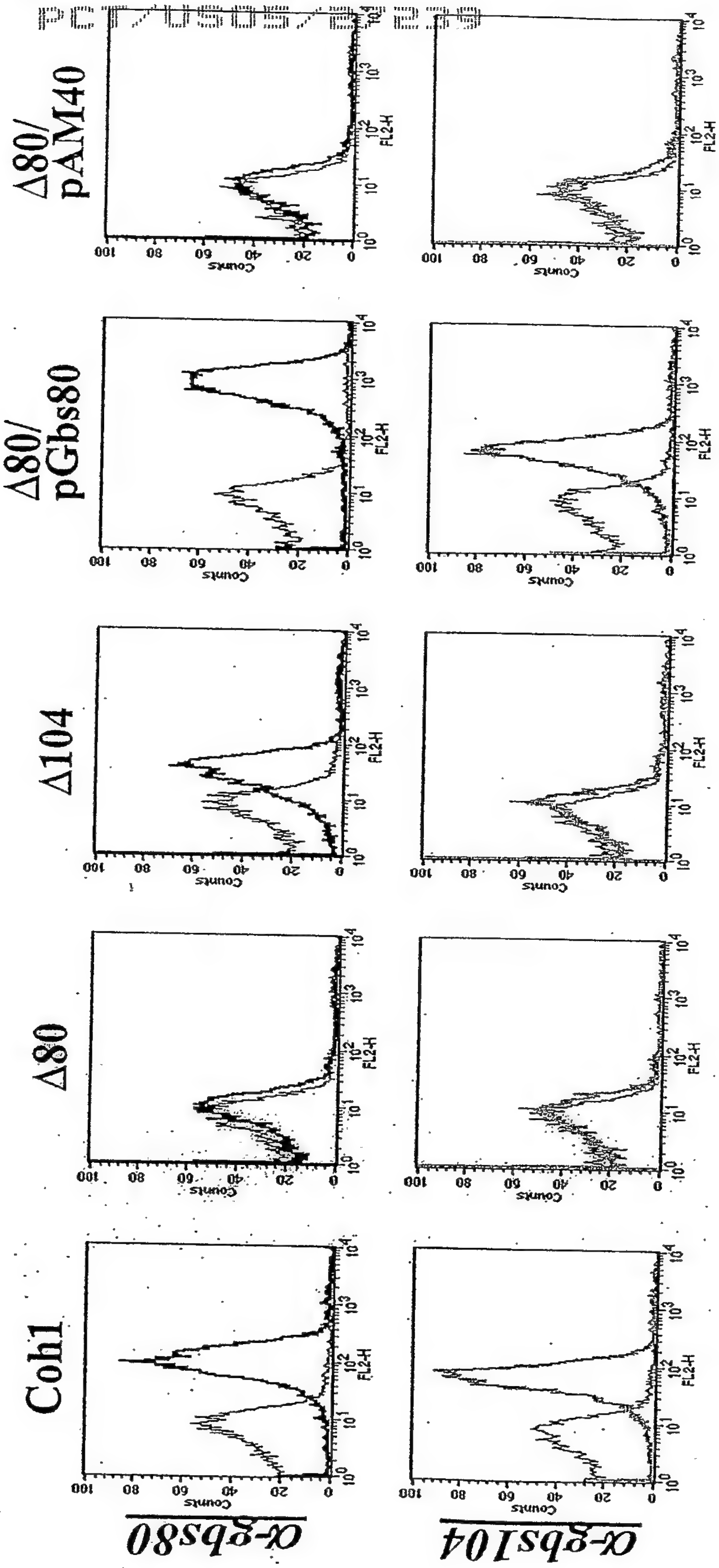


Figure 8: gbs80 is required for surface localization of gbs104

Figure 9: sortases 647 & 648 play a semi-redundant role in surface exposure of gbs80 and gbs104

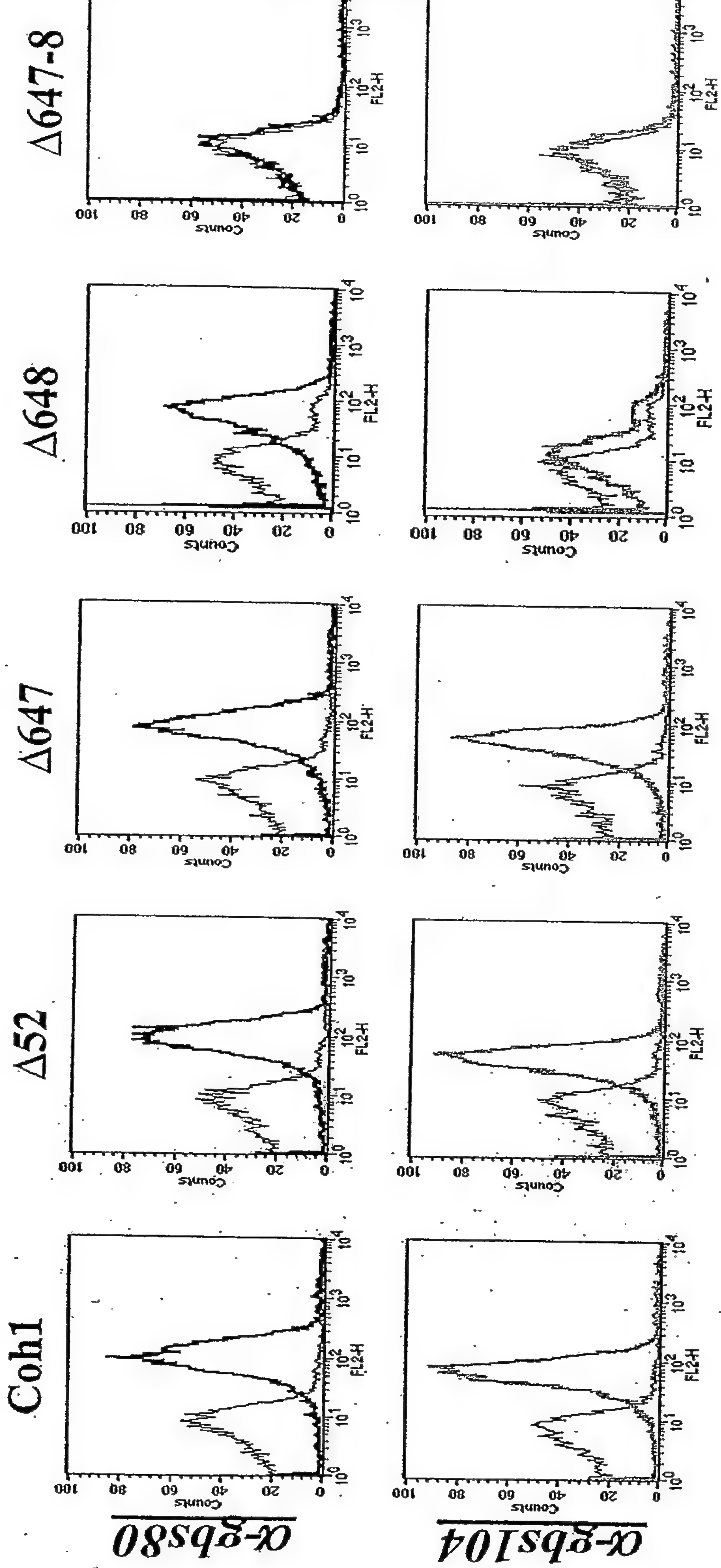
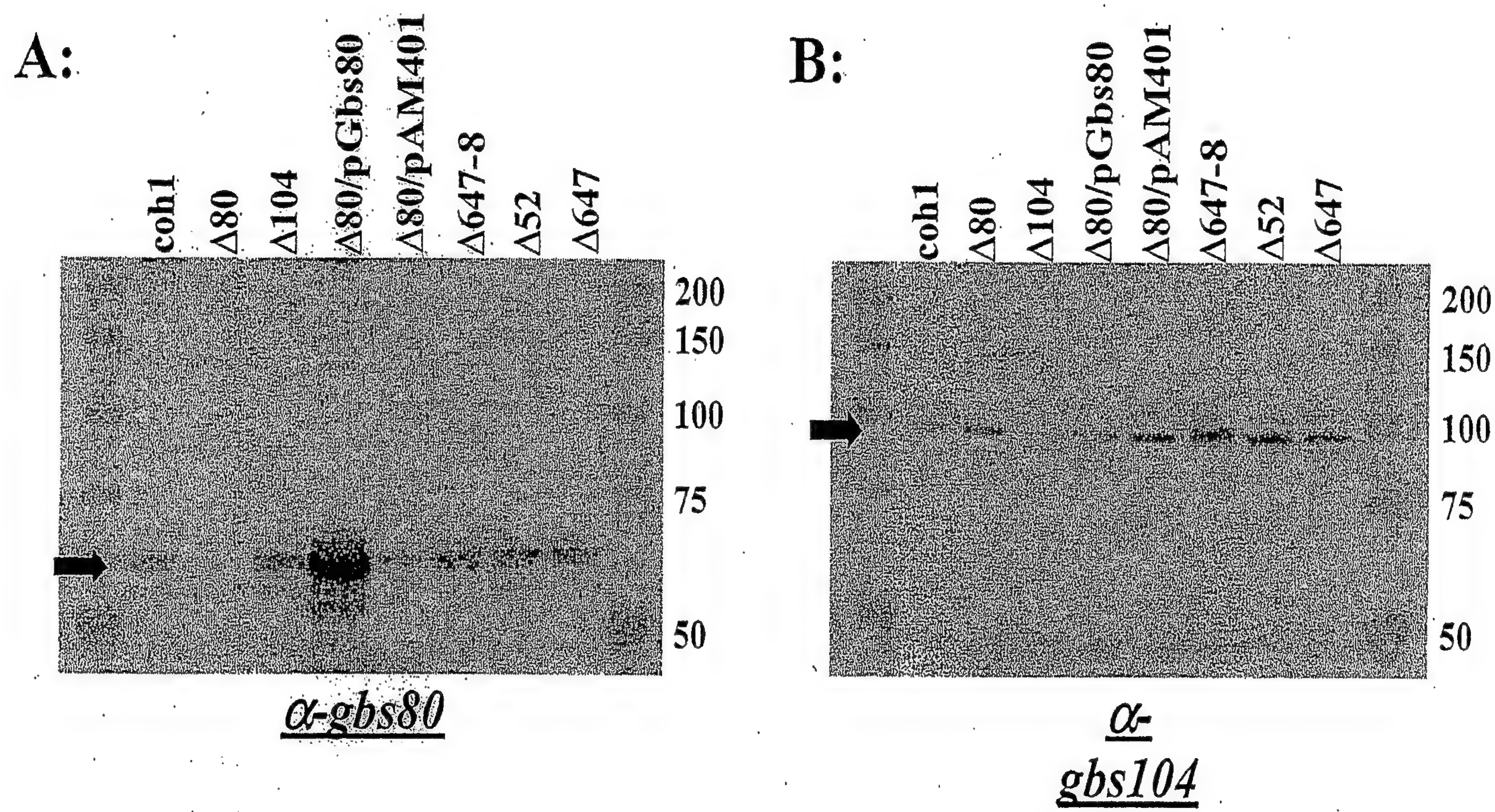


Figure 10: Western blot of mutant strains



PCT/US05/27239

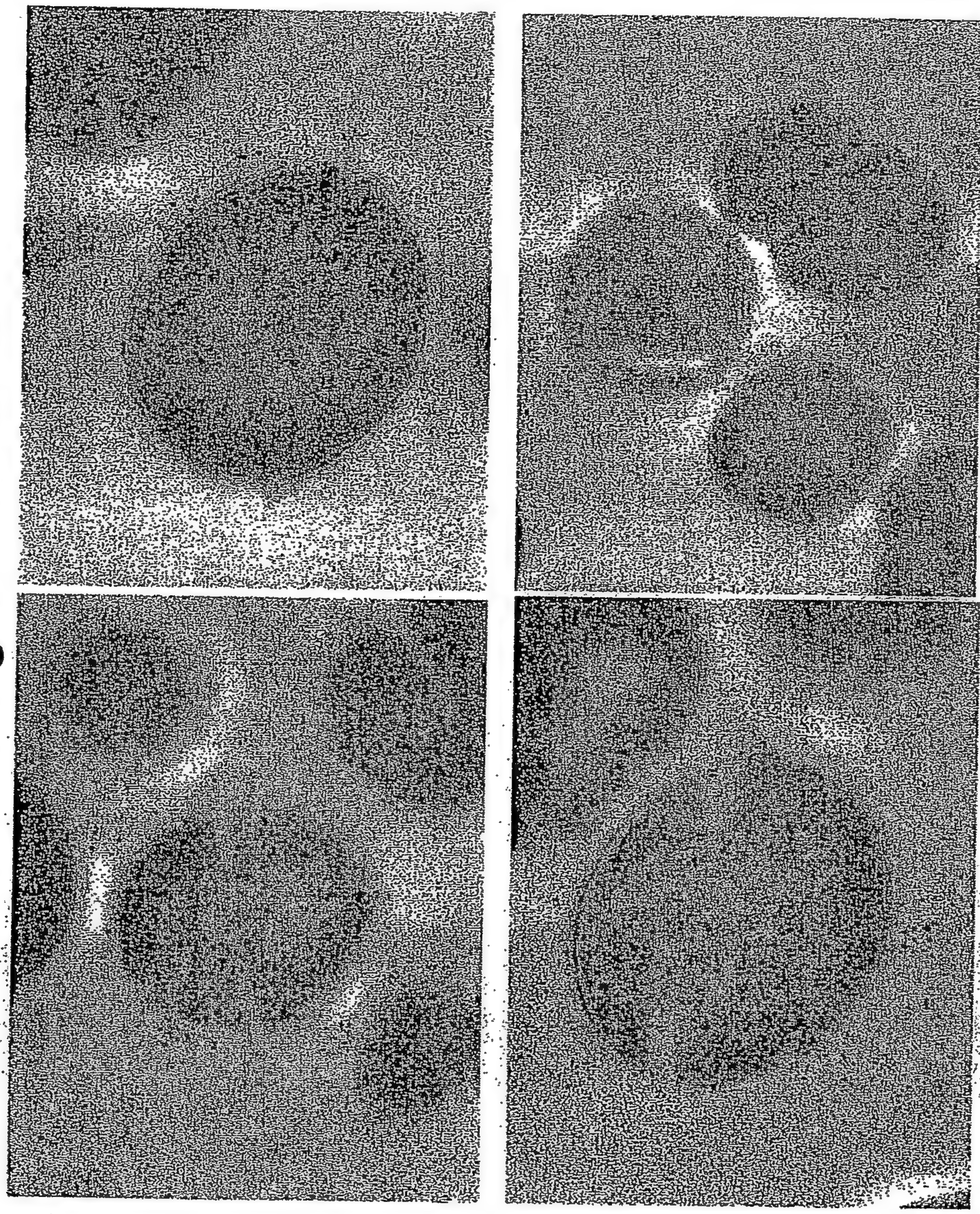
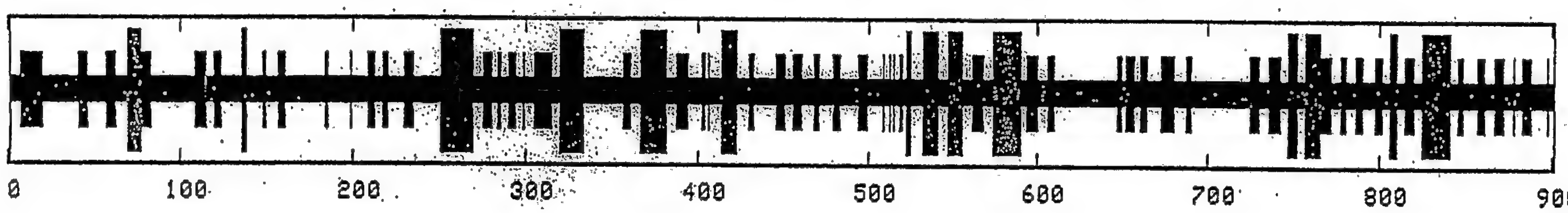


FIGURE 11: Pre-embedding IEM staining of GBS 80

FIGURE 12: Predicted Secondary Structure for GBS 067

PHD SECONDARY STRUCTURE PREDICTION for GBS 067

```
         10        20        30        40        50        60        70
          |         |         |         |         |         |         |
MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQNKPLSKATFVLKTTAHPESK
CccccceeeeeEEeceeeEcccccccccccCCCCCcceEEEEeCCCCCCCCccEEEEEecCCCcch
IEKVTAELTGEATFDNLIPGDYTLSEETAPEGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELD
hHHhhheeEeeeccCCCCCCccCcCCCCcccccccCeEEEEEEeCCcEEEEeCCCCCCCCCchHHhc
KQYPPTGIYEDTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEVGDLAHNKYKI
CCCCCCcceecCccccceeEEecccccCCCCcccccCCCCCCcccceeccCCCccccccccecccccccCCcEE
ELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQRHNKAKKAAEALGTAVKDILGANSDNRVALV
EEEccCcEEEecCCCCCCeEEEEEcCCCCCCCCCCCCChhHHHHHHHHHHHHHHHHhcCcCCeEEEE
TYGSDIFDGRSVDVVKGFKEDDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTNGL
EcccceecCCCeeEEecccecCCCcceeeEEEEEEecCCcchhHHHHHHHHHHHhcCCCCCCCCCCCCC
TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTRSYAINNFKLGASYESQFEQM
CcccccceEEEecccccchhHHHHHHHHHHHHHHcccCCeEEEEEeCCCCCccceecceccCCCCchHHHHHH
KKNGYLNKSNFLLTDKPEDIKGNGESYFLFPLDSYQTQIISGNLQKLHYLDLNLNYPKGTIYRNGPVKEH
HHhCCcccceeeecCCCCecCCCCceEEeeecCcCeeEEEecCcccccceeecCCCCCCeEEeCCCcCCC
GTPTKLYINSLKQKNYDIFNFGIDISGFRQVYNEEYKKNQDGTFQKLKEEAFKLSDGEITELMRSFSSKP
CCCceEeEEcccccccccecccececcceecchHHhcccCCcHHHHHHHHHhcccCchHHHHHHHHcCCC
EYYTPIVTSADTSNNEILSKIQQQFETILTKENSIVNGTIEDPMGDKINLQLGNGQTLQPSDYTLQGNDG
cceeEEEeecCCCcHHHHHHHHHHHHHHhhcccceeeeeEecCCCCeeEEecCCCcccccCCCCcccCCCC
SVMKDGIATGGPNNDGGILKGVKLEYIGNKLYVRGLNLGEGQKVTLTYDVKLDDSFISNKFYDTNGRTTL
ccccccCCCCCCCCCCeeeeeEEEEeCCcEEEEEccCCCCCEEEEEEEeecCCCcceeeecCCCCCccC
NPKSEDPNTLRDFPIPKIRDVREYPTITIKNEKKLGEIEFIKVDKDNNKLLLKGATFELQEFNEDYKLYL
CCCCCCCCCCCCCCCCCCcccccCCCeEEEEeCCCCCeEEEEEeCCCChhhHHccCchHHHHhchhhceee
PIKNNNSKVVTGENGKISYKDLKDGKYQLIEAVSPEDYQKITNKPILTFEVVKGSIKNIIAVNKQISEYH
eecCCCCcEEEccCCcEeEEecccCCcEEEEECCCchhHhhCCcCeEEEEeccccchhhhhhhHHHHHHH
EEGDKHLITNTHIPPKGIIPMTGGKGILSFILIGGAMMSIAGGIYIWKRYKKSSDMSIKKD
HhCCceEEeCCCCCCCCeEeecCCCccEEEEecccccceccCCeeEEEccCCCCCCCcecCC
Sequence length :   901
PHD :
  Alpha helix      (Hh) :   148 is  16.43%    3₁₀ helix   (Gg) :    0 is   0.00%
  Pi helix         (Ii) :     0 is   0.00%    Beta bridge (Bb) :    0 is   0.00%
  Extended strand  (Ee) :   243 is  26.97%    Beta turn   (Tt) :    0 is   0.00%
  Bend region      (Ss) :     0 is   0.00%    Random coil (Cc) :  510 is  56.60%
  Ambigous states  (?)  :     0 is   0.00%    Other states     :    0 is   0.00%
```



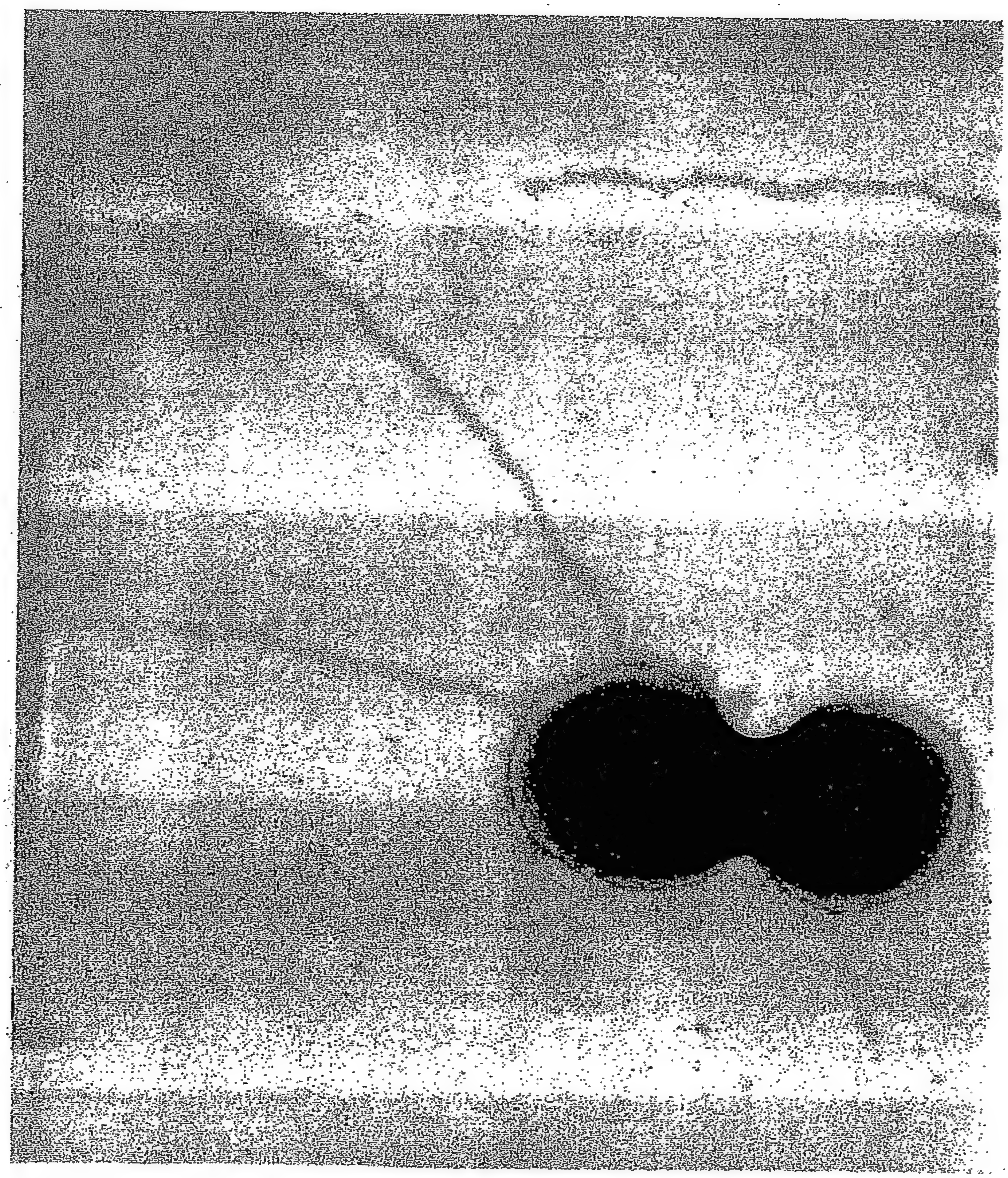


Figure 13

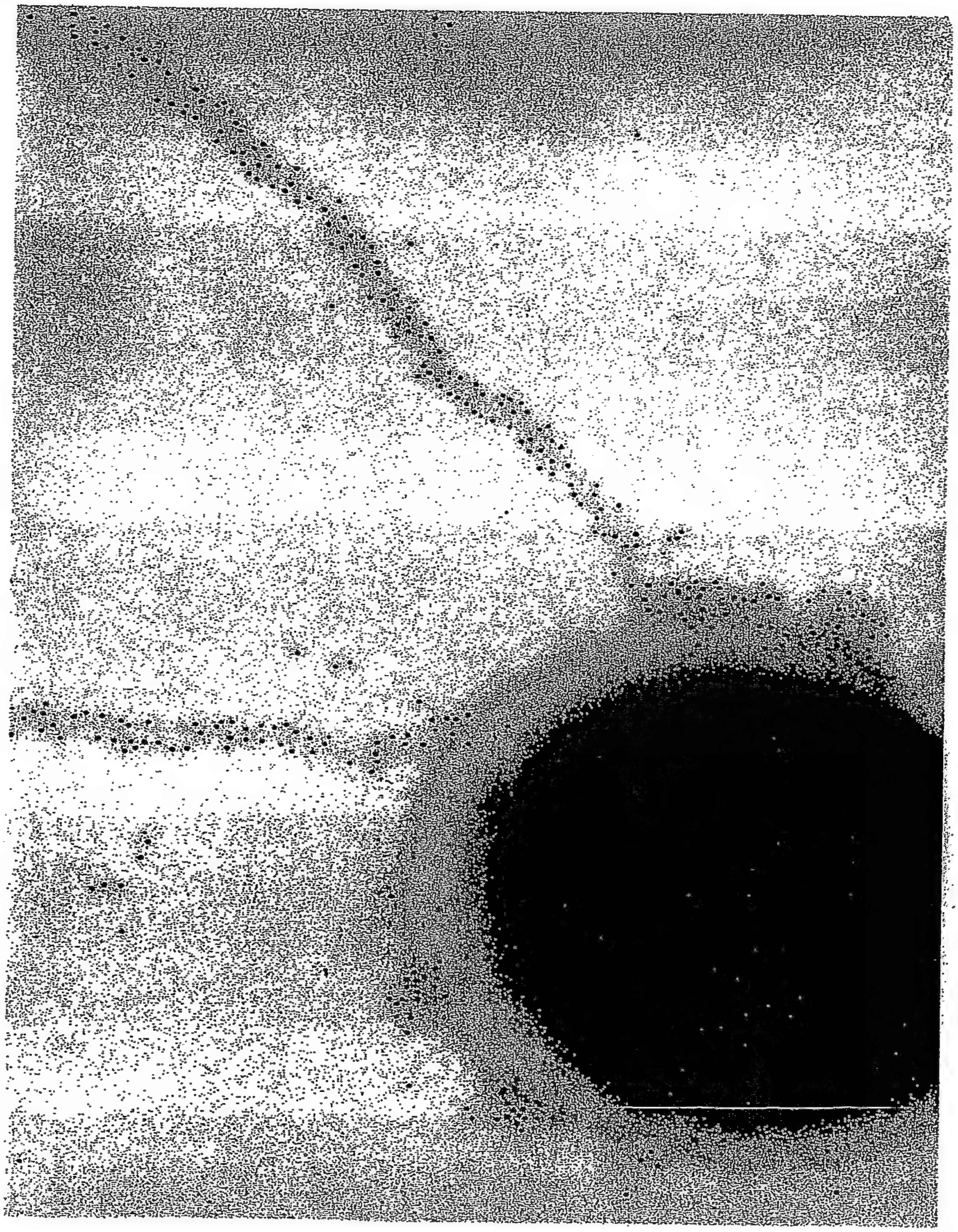


Figure 14

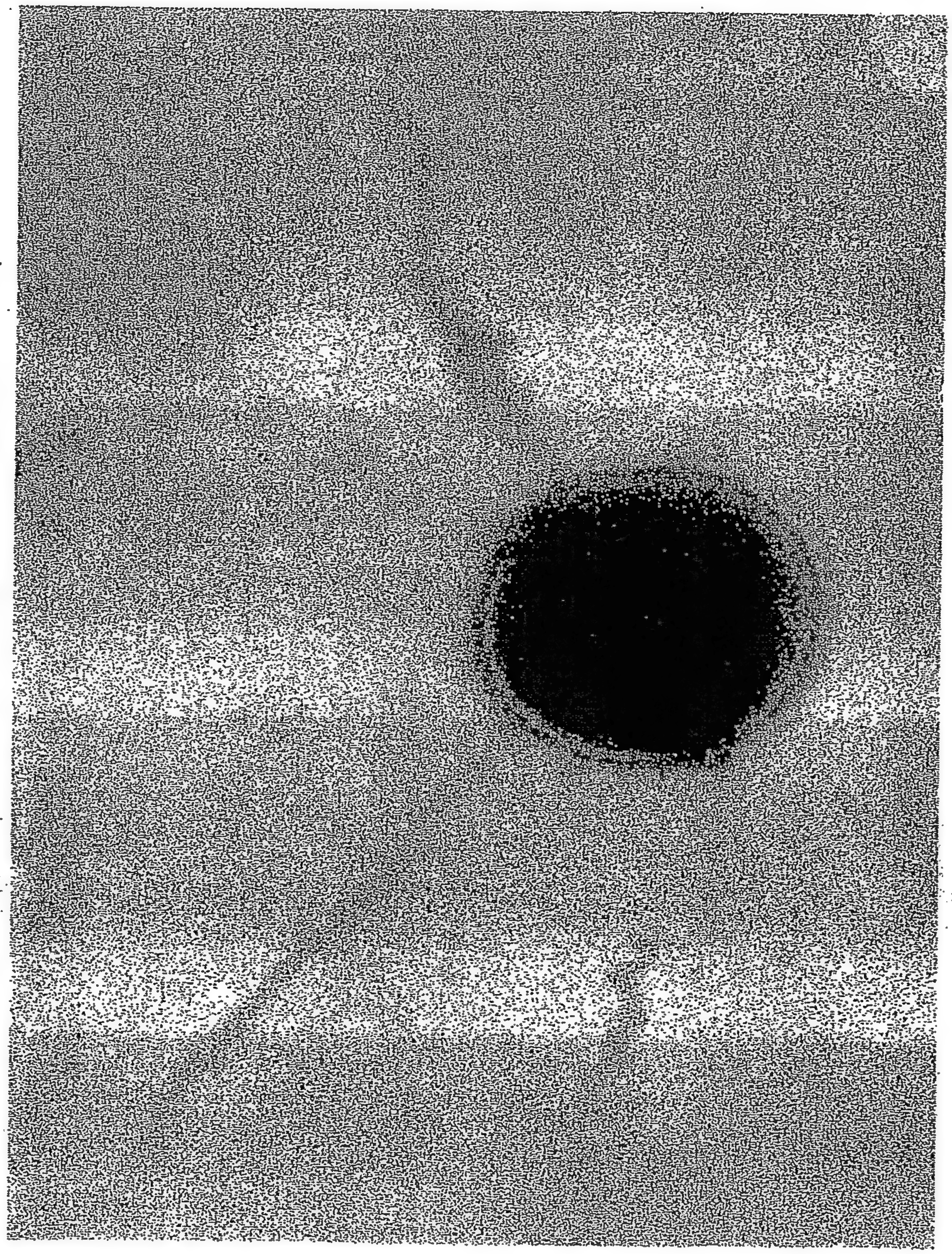


Figure 15

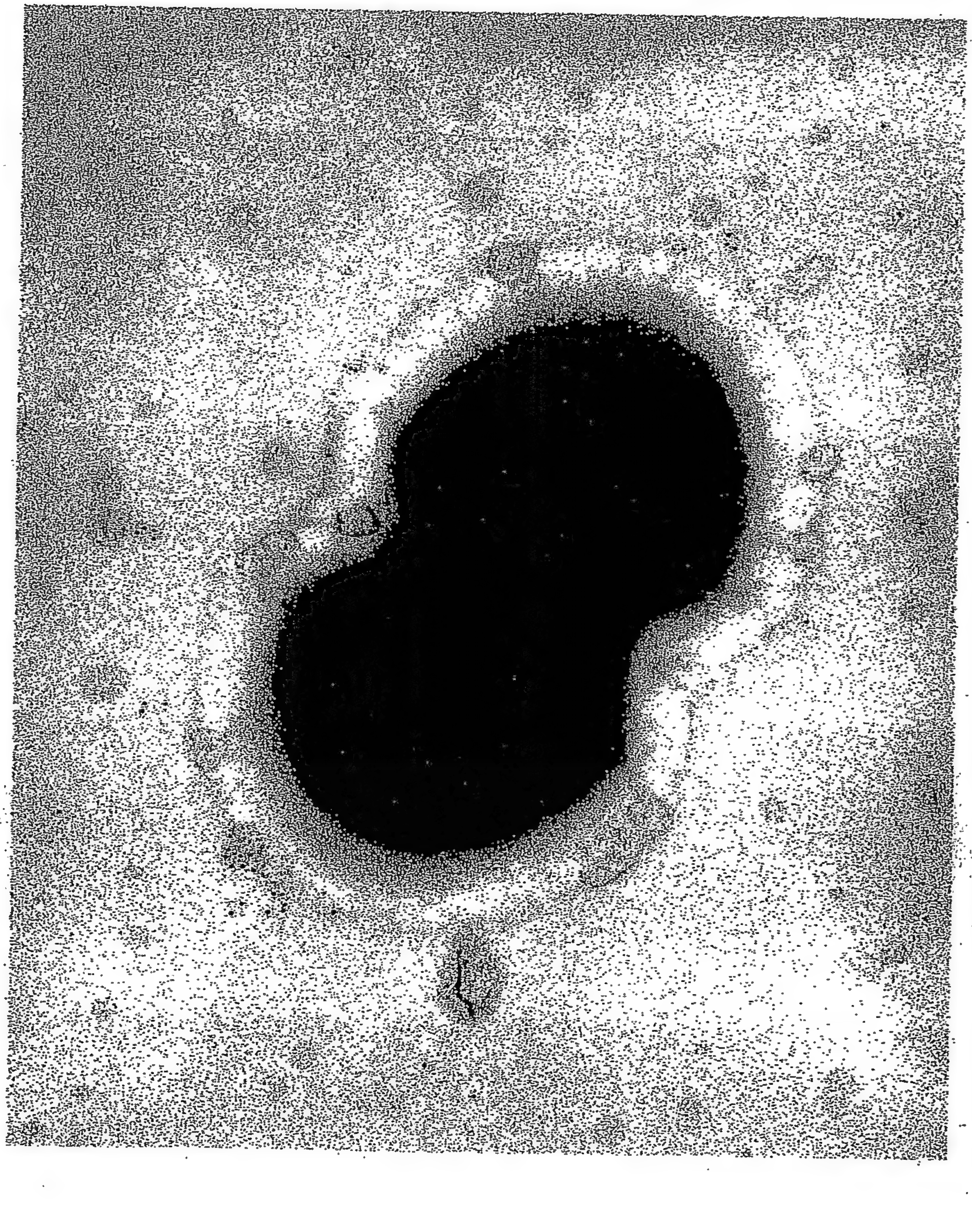


Figure 16

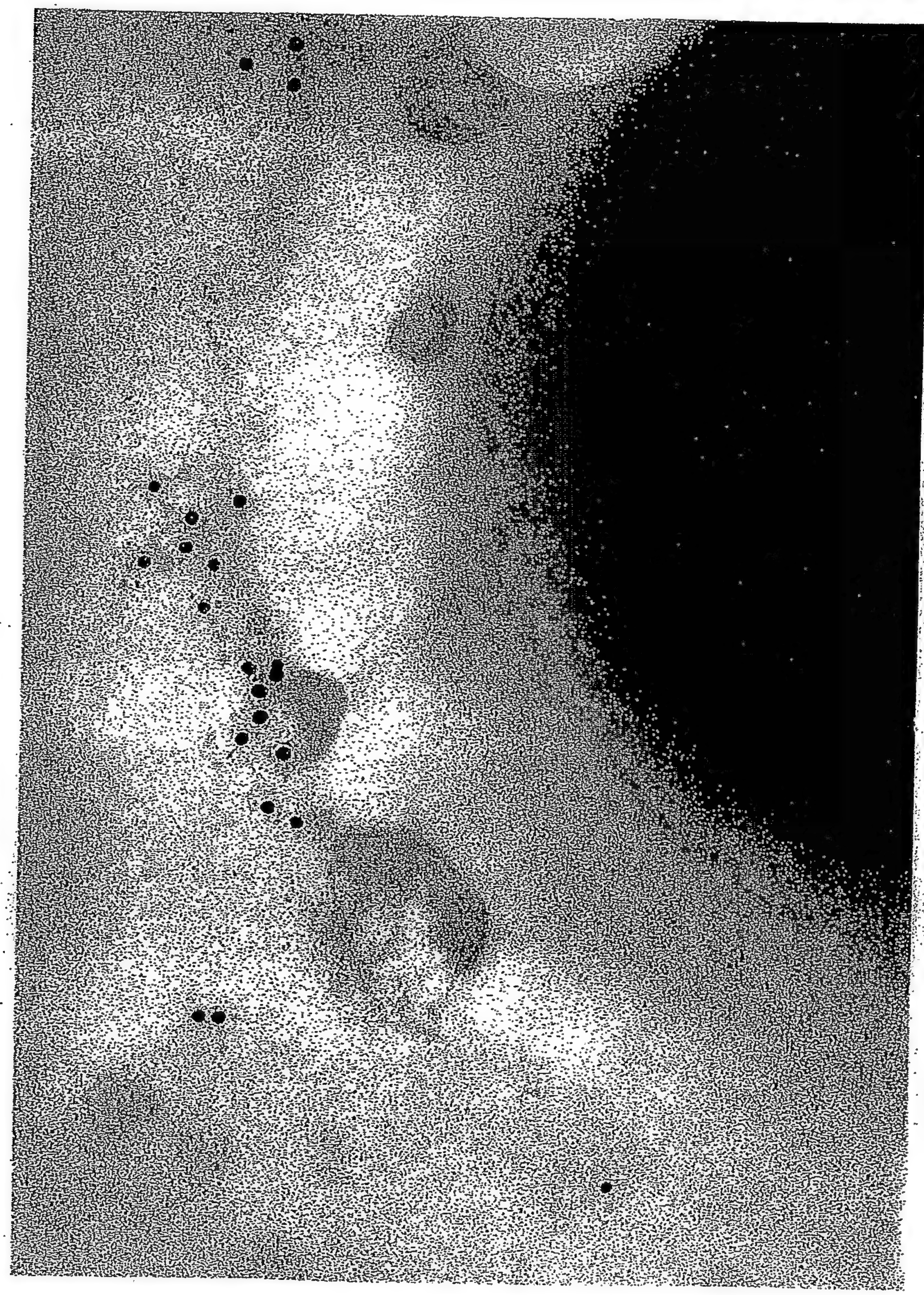


Figure 17

```
     TAACACGGATTGCAAATAACATCAATCAAATTGCGCGTGCTATCAACCAA Majority
              10        20        30        40        50
1    TAACACGGATTGCAAATAACATCAATCAAATTGCGCGTGCTATCAACCAA 2603_ai1.seq
1    -------------------------------------------------- 18rs21_ai1.seq
1    TAACACGGATTGCAAATAACATCAATCAAATTGCGCGTGCTATCAACCAA coh1_ai1.seq
1    TAACACGGATTGCAAATAACATCAATCAAATTTCGCGTGCTATCAACCAA cjb111_ai1.seq
1    TAACACGGATTGCAAATAACATCAATCAAATTGCGCGTGCTATCAACCAA nem316_ai1.seq
1    TAACACGGATTGCAAATAACATCAATCAAATTTCGCGTGCTATCAACCAA a909_ai1.seq

     AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA Majority
              60        70        80        90       100
51   AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA 2603_ai1.seq
1    -------------------------------------------------- 18rs21_ai1.seq
51   AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA coh1_ai1.seq
51   AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA cjb111_ai1.seq
51   AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA nem316_ai1.seq
51   AGCCATCTAATTTCTCAGGCTCAGTATCACTTTCATCGAGAAGCACGTTA a909_ai1.seq

     AGCCCAACAGCAGGGGGGGTTGCTCTGATTCGGTTAAGTAATAAGCAAAA Majority
             110       120       130       140       150
101  AGCCCAACAGCAGGGGGG-TTGCTCTGATTCGGTTAAGTAATAAGCAAAA 2603_ai1.seq
1    -------------------------------------------------- 18rs21_ai1.seq
101  AGCCCAACAGCAGGGGGGGTTGCTCTGATTCGGTTAAGTAATAAGCAAAA coh1_ai1.seq
101  AGCCCAACAGCAGGGGGGG-TGCTCTGATTCGGTTAAGTAATAAGCAAAA cjb111_ai1.seq
101  AGCCCAACAGCAGGGGGGGGTGCTCTGATTCGGTTAAGTAATAAGCAAAA nem316_ai1.seq
101  AGCCCAACAGCAGGGGGGGTTGCTCTGATTCGGTTAAGTAATAAGCAAAA a909_ai1.seq

     TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT Majority
             160       170       180       190       200
150  TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT 2603_ai1.seq
1    -------------------------------------------------- 18rs21_ai1.seq
151  TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT coh1_ai1.seq
150  TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT cjb111_ai1.seq
51   TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT nem316_ai1.seq
51   TCCTCACCAATTTTCCCAGTAATAAGAGGCGTTGTTGAGGTATAGGGATT a909_ai1.seq

     ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG Majority
             210       220       230       240       250
00   ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG 2603_ai1.seq
     --------------------------------GACCATTCCCCATAAAAG 18rs21_ai1.seq
01   ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG coh1_ai1.seq
00   ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG cjb111_ai1.seq
01   ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG nem316_ai1.seq
01   ACCTGTGCCATAGTCGATGATGGTCACAAAATGACCATTCCCCATAAAAG a909_ai1.seq

     GGCCAACCAAAACTTCACCTGTGGCCGTTTTTTGATGTCAACACCAGTAT Majority
             260       270       280       290       300
50   GGCCAACCAAAACTTCACCTGTGACCGTTTTTTGATGTCAACACCAGTAT 2603_ai1.seq
9    GGCCAACCAAAACTTCACCTGTGACCGTTTTTTGATGTCAACACCAGTAT 18rs21_ai1.seq
51   GGCCAACCAAAACTTCACCTGTGGCCGTTTTTTGATGTCAACACCAGTAT coh1_ai1.seq
50   GGCCAACCAAAACTTCACCTGTGGCCGTTTTTTGATGTCAACACCAGTAT cjb111_ai1.seq
51   GGCCAACCAAAACTTCACCTGTGGCCGTTTTTTGATGTCAACACCAGTAT nem316_ai1.seq
51   GGCCAACCAAAACTTCACCTGTGGCCGTTTTTTGATGTCAACACCAGTAT a909_ai1.seq

     TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA Majority
             310       320       330       340       350
00   TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA 2603_ai1.seq
9    TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA 18rs21_ai1.seq
01   TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA coh1_ai1.seq
00   TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA cjb111_ai1.seq
01   TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA nem316_ai1.seq
01   TTTGGATGTACCCCTTAACATGCCCTTTGGTATCTGCAACAGAGATAATA a909_ai1.seq
```

Figure 18

```
     TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC Majority
              360       370       380       390       400
50   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC 2603_ai1.seq
19   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC 18rs21_ai1.seq
51   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC coh1_ai1.seq
50   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC cjb111_ai1.seq
51   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC nem316_ai1.seq
51   TGACCAAAGGAAGAATCACCGATGACTTTAACCGTAATCTTGCTATCGCC a909_ai1.seq

     TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG Majority
              410       420       430       440       450
00   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG 2603_ai1.seq
69   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG 18rs21_ai1.seq
01   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG coh1_ai1.seq
00   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG cjb111_ai1.seq
01   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG nem316_ai1.seq
01   TTTTTGATTAGCCGCTAATATTTGATTAGCAATCAGGGTGCGACCAAGAG a909_ai1.seq

     CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA Majority
              460       470       480       490       500
50   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA 2603_ai1.seq
19   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA 18rs21_ai1.seq
51   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA coh1_ai1.seq
50   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA cjb111_ai1.seq
51   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA nem316_ai1.seq
51   CAACTGTTGAAGACGACAAGGTATGATGTTTTTCCTGAGCCAATTTAACA a909_ai1.seq

     GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT Majority
              510       520       530       540       550
00   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT 2603_ai1.seq
69   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT 18rs21_ai1.seq
01   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT coh1_ai1.seq
00   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT cjb111_ai1.seq
01   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT nem316_ai1.seq
01   GTCTCGGTGCTATCGAGGACATAAGACCTGAAGGCACCAGAGGCAGAAAT a909_ai1.seq

     TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT Majority
              560       570       580       590       600
50   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT 2603_ai1.seq
19   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT 18rs21_ai1.seq
51   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT coh1_ai1.seq
50   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT cjb111_ai1.seq
51   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT nem316_ai1.seq
51   TGATTTAATAATTTTATCCATAACCTATGTTATAGCACAAAGAGAGAGTT a909_ai1.seq

     TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG Majority
              610       620       630       640       650
00   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG 2603_ai1.seq
69   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG 18rs21_ai1.seq
01   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG coh1_ai1.seq
00   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG cjb111_ai1.seq
01   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG nem316_ai1.seq
01   TGTATGGACTGGATTAGCTGAAAATTTTAGACTAAAAAGTAGTGCCAGTG a909_ai1.seq

     GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTGTTT Majority
              660       670       680       690       700
50   GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTCTTT 2603_ai1.seq
19   GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTCTTT 18rs21_ai1.seq
51   GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTGTTT coh1_ai1.seq
50   GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTGTTT cjb111_ai1.seq
51   GATTTTGAAAGGTAACAAAAGAGTCTTCCGTTTAAACGAGAAGGCTGTTT nem316_ai1.seq
51   GATTTTGAAAGGTAACAAAAGAGTTTTCCGTTTAAACGAGAAGGCTGTTT a909_ai1.seq
```

FIGURE 18 A

Alignment Report of AI-1_alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM Page 3

```
      TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT Majority
                710       720       730       740       750
'00   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT 2603_all.seq
:69   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT 18rs21_all.seq
'01   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT coh1_all.seq
'00   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT cjb111_all.seq
'01   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT nem316_all.seq
'01   TTAGTCGTTTCAGATGAAGGCAAAAGGACGATGAAATAATTCCGTACCTT a909_all.seq

      CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA Majority
                760       770       780       790       800
50    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA 2603_all.seq
19    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA 18rs21_all.seq
51    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA coh1_all.seq
50    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA cjb111_all.seq
51    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA nem316_all.seq
51    CATGGATTGCTATGTTACTGGCATGAGGTCTCACGATATTTAGTAAGATA a909_all.seq

      TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG Majority
                810       820       830       840       850
00    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG 2603_all.seq
69    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG 18rs21_all.seq
01    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG coh1_all.seq
00    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG cjb111_all.seq
01    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG nem316_all.seq
01    TTCGTTTGAAGATATTCCCACGTATTTTTTAAAGGTTTTAAGAAAATGTG a909_all.seq

      TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT Majority
                860       870       880       890       900
50    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT 2603_all.seq
19    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT 18rs21_all.seq
51    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT coh1_all.seq
50    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT cjb111_all.seq
51    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT nem316_all.seq
51    TGGTGTCGTAAAAATGTAATAATTTCGCTACTTCCCCAATCGGTACCCCT a909_all.seq

      CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT Majority
                910       920       930       940       950
00    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT 2603_all.seq
69    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT 18rs21_all.seq
01    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT coh1_all.seq
00    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT cjb111_all.seq
01    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT nem316_all.seq
01    CTTTGCAGTAGTAATTGTCCCTCCTTAATTTTTGCCTTTAGAATATAACT a909_all.seq

      TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACTGAGCGTAAGTGGGAAT Majority
                960       970       980       990       1000
50    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACTGAGCGTAAGTGGGAAT 2603_all.seq
19    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACTGAGCGTAAGTGGGAAT 18rs21_all.seq
51    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACTGAGCGTAAGTGGGAAT coh1_all.seq
50    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACCGAGCGTAAGTGGGAAT cjb111_all.seq
51    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACCGAGCGTAAGTGGGAAT nem316_all.seq
51    TTGCAAGGAAATGTCAGAGTATTTTTTAAAAACCGAGCGTAAGTGGGAAT a909_all.seq

      CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG Majority
                1010      1020      1030      1040      1050
1000  CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG 2603_all.seq
9     CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG 18rs21_all.seq
1001  CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG coh1_all.seq
1000  TTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG cjb111_all.seq
1001  CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG nem316_all.seq
1001  CTGAGATATATAGGTAGTTGGCAATATCAGATACTTTGAGTTTGGAGTAG a909_all.seq
```

FIGURE 18 B

```
     AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC Majority
              1060      1070      1080      1090      1100
1050 AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC 2603_ail.seq
819  AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC 18rs21_ail.seq
1051 AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC coh1_ail.seq
1050 AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC cjb111_ail.seq
1051 AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC nem316_ail.seq
1051 AGGTCGTTGTCCACATAATGGAGAATACTATTGTACATTTGCTGCTTGTC a909_ail.seq

     AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC Majority
              1110      1120      1130      1140      1150
1100 AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC 2603_ail.seq
869  AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC 18rs21_ail.seq
1101 AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC coh1_ail.seq
1100 AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC cjb111_ail.seq
1101 AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC nem316_ail.seq
1101 AGAGATGCTCTTATTGGTTAAGGATTCTGAAAAATCAATAAGAGCTGCAC a909_ail.seq

     AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC Majority
              1160      1170      1180      1190      1200
1150 AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC 2603_ail.seq
919  AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC 18rs21_ail.seq
1151 AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC coh1_ail.seq
1150 AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC cjb111_ail.seq
1151 AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTATTCGTTTAAAGCC nem316_ail.seq
1151 AGCGAATTCTTGAAACATCAATAAGATCAGGAGCCTCTTCGTTTAAAGCC a909_ail.seq

     ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC Majority
              1210      1220      1230      1240      1250
1200 ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC 2603_ail.seq
969  ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC 18rs21_ail.seq
1201 ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC coh1_ail.seq
1200 ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTTCTC cjb111_ail.seq
1201 ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC nem316_ail.seq
1201 ATATAGTGCTTTACCAGCGCATAACTTTTAGCCACATCAGTATTTTCCTC a909_ail.seq

     GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT Majority
              1260      1270      1280      1290      1300
250  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT 2603_ail.seq
019  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT 18rs21_ail.seq
251  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT coh1_ail.seq
250  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT cjb111_ail.seq
251  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT nem316_ail.seq
251  GAAACTTAATTCTAGTAATTTTGTTAAGTAAACAACAGTTAAGTTCTTTT a909_ail.seq

     CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG Majority
              1310      1320      1330      1340      1350
300  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG 2603_ail.seq
069  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG 18rs21_ail.seq
301  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG coh1_ail.seq
300  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG cjb111_ail.seq
301  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG nem316_ail.seq
301  CAGCTCTTAGGGCAGGGATTGAAGATGAGGTAACACTGGATGATGGGAGG a909_ail.seq

     CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC Majority
              1360      1370      1380      1390      1400
350  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC 2603_ail.seq
119  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC 18rs21_ail.seq
351  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC coh1_ail.seq
350  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC cjb111_ail.seq
351  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC nem316_ail.seq
351  CGATTAATTTCTTGCTTTAACAGTTGAGTGTTACCCAGCTTAACGAGATC a909_ail.seq
```

FIGURE 18 C

AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT Majority
1410 1420 1430 1440 1450
400 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT 2603_ai1.seq
169 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT 18rs21_ai1.seq
401 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT coh1_ai1.seq
400 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT cjb111_ai1.seq
401 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT nem316_ai1.seq
401 AATAATGTGATTGAGATGCTTTAAAACAGTGGGTAACTGAAAAGAGTTTT a909_ai1.seq

TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT Majority
1460 1470 1480 1490 1500
450 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT 2603_ai1.seq
219 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT 18rs21_ai1.seq
451 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT coh1_ai1.seq
450 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT cjb111_ai1.seq
451 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT nem316_ai1.seq
451 TCTTAGTATGTTTTAGGTGAAGAACAATATCAGGATCCGCAACAATCTGT a909_ai1.seq

TCTGACTCTTCTAATAAATGATTGATGGCTTGTTGGCAACTAGCCTCAAA Majority
1510 1520 1530 1540 1550
500 TCTGACTCTTCTAATAAATGATTGATGACTTGTTGGCAACTAGCCTCAAA 2603_ai1.seq
269 TCTGACTCTTCTAATAAATGATTGATGACTTGTTGGCAACTAGCCTCAAA 18rs21_ai1.seq
501 TCTGACTCTTCTAATAAATGATTGATGACTTGTTGGCAACTAGCCTCAAA coh1_ai1.seq
500 TCTGACTCTTCTAATAAATGATTGATGGCTTGTTGGCAACTAGCCTCAAA cjb111_ai1.seq
501 TCTGACTCTTCTAATAAATGATTGATGGCTTGTTGGCAACTAGCCTCAAA nem316_ai1.seq
501 TCTGACTCTTCTAATAAATGATTGATGGCTTGTTGGCAACTAGCCTCAAA a909_ai1.seq

CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG Majority
1560 1570 1580 1590 1600
550 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG 2603_ai1.seq
319 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG 18rs21_ai1.seq
551 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG coh1_ai1.seq
550 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG cjb111_ai1.seq
551 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG nem316_ai1.seq
551 CTGTGTTTGGAAAAAGGCATCGATAGACACAAGAAGACTACGTATACTGG a909_ai1.seq

TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA Majority
1610 1620 1630 1640 1650
600 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA 2603_ai1.seq
369 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA 18rs21_ai1.seq
601 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA coh1_ai1.seq
600 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA cjb111_ai1.seq
601 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA nem316_ai1.seq
601 TAGTAGGAAAACAAGGGACAAGCTTTATATAGGATAAGATTTCTTTTTTA a909_ai1.seq

CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT Majority
1660 1670 1680 1690 1700
650 TTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT 2603_ai1.seq
419 CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT 18rs21_ai1.seq
651 CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT coh1_ai1.seq
650 CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT cjb111_ai1.seq
651 CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT nem316_ai1.seq
651 CTACGATGAGAAAATTGTTCTAGAAAGCGACTGGATAACTGTTCTTGCCT a909_ai1.seq

ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT Majority
1710 1720 1730 1740 1750
700 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT 2603_ai1.seq
469 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT 18rs21_ai1.seq
701 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT coh1_ai1.seq
700 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT cjb111_ai1.seq
701 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT nem316_ai1.seq
701 ATTGATATCAGGGCTATAGGGATAAAATGGTCCAATAGCAATAAGATATT a909_ai1.seq

FIGURE 18 D

```
     GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA Majority
              1760      1770      1780      1790      1800
750  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA 2603_ai1.seq
519  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA 18rs21_ai1.seq
751  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA coh1_ai1.seq
750  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA cjb111_ai1.seq
751  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGACCCTCATAAACA nem316_ai1.seq
751  GACAGACAGGAAAAATTAAGAATGATTCTTCAAAAAGATCCTCATAAACA a909_ai1.seq

     GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA Majority
              1810      1820      1830      1840      1850
800  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA 2603_ai1.seq
569  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA 18rs21_ai1.seq
801  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA coh1_ai1.seq
800  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA cjb111_ai1.seq
801  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA nem316_ai1.seq
801  GTGATATCTTGGTTATAAGGGATAGCTAAATGTTTTAAAAACTGATAGTA a909_ai1.seq

     AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC Majority
              1860      1870      1880      1890      1900
850  AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC 2603_ai1.seq
619  AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC 18rs21_ai1.seq
851  AGGCAACAGATAGTCTTCGTTACTATATAACTGAACGAGTTCCTTGTCTC coh1_ai1.seq
850  AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC cjb111_ai1.seq
851  AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC nem316_ai1.seq
851  AGGCAACAGATAGTCTTCGTTACCATATAACTGAACGAGTTCCTTGTCTC a909_ai1.seq

     GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA Majority
              1910      1920      1930      1940      1950
900  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA 2603_ai1.seq
669  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA 18rs21_ai1.seq
901  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA coh1_ai1.seq
900  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA cjb111_ai1.seq
901  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA nem316_ai1.seq
901  GTGACATGACTGAAATAGGTAGTTGAGATATGGTATGCAATGTTTGAACA a909_ai1.seq

     TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC Majority
              1960      1970      1980      1990      2000
950  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC 2603_ai1.seq
719  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC 18rs21_ai1.seq
951  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC coh1_ai1.seq
950  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC cjb111_ai1.seq
951  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC nem316_ai1.seq
951  TGTTTAAAATCGAATGTAACCATTTGATAGACCGCCTTCATTATCATTTC a909_ai1.seq

     TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC Majority
              2010      2020      2030      2040      2050
000  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC 2603_ai1.seq
769  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC 18rs21_ai1.seq
001  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC coh1_ai1.seq
000  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC cjb111_ai1.seq
001  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC nem316_ai1.seq
001  TAGAATTTTTCTTTAGGTTTGTAAAGACTACAAAATAAAATGATGAAAAC a909_ai1.seq

     AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC Majority
              2060      2070      2080      2090      2100
050  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC 2603_ai1.seq
819  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC 18rs21_ai1.seq
051  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC coh1_ai1.seq
050  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC cjb111_ai1.seq
051  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC nem316_ai1.seq
051  AACTATCTTGTGGATACACTAAAAAGACACGCTAATTAGCAAACTCTCTC a909_ai1.seq
```

FIGURE 18 E

```
      TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG Majority
               2110      2120      2130      2140      2150
2100  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG 2603_all.seq
1869  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG 18rs21_all.seq
2101  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG coh1_all.seq
2100  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG cjb111_all.seq
2101  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG nem316_all.seq
2101  TTCATCATCTCTCACCATTATTATACTACTATTTATATGACAAATAAAGG a909_all.seq

      TGATTTTGTTAAAAATATAACTTTGAAAATCCACATATATTTTTAATCTT Majority
               2160      2170      2180      2190      2200
2150  TGATTTTGTTAAAA-TATAACTTTGAAAATCCACATATATTTTTAATCTT 2603_all.seq
1919  TGATTTTGTTAAAA-TATAACTTTGAAAATCCACATATATTTTTAATCTT 18rs21_all.seq
2151  TTATTTTGTTAAAA-TATAACTTTGAAAATCCACATATATTTTTAATCTT coh1_all.seq
2150  TGATTTTGTTAAAAATATAACTTTGAAAATCCACATATATTTTTAATCTT cjb111_all.seq
2151  TGATTTTGTTAAAAATATAACTTTGAAAATCCACATATATTTTTAATCTT nem316_all.seq
2151  TGATTTTGTTAAAAATATAACTTTGAAAATCCACATATATTTTTAATCTT a909_all.seq

      CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA Majority
               2210      2220      2230      2240      2250
2199  CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA 2603_all.seq
1968  CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA 18rs21_all.seq
2200  CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA coh1_all.seq
2200  CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA cjb111_all.seq
2201  CCGTCTGAAAAAAATAAATAAAAATAGTAAAAATAAACACGAATTTAAAA nem316_all.seq
2201  CCGTCTGAAAAAA-TAAATAAAAATAGTAAAAATAAACACGAATTTAAAA a909_all.seq

      TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT Majority
               2260      2270      2280      2290      2300
2248  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT 2603_all.seq
2017  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT 18rs21_all.seq
2249  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT coh1_all.seq
2249  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT cjb111_all.seq
2251  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT nem316_all.seq
2250  TAAGCAAATTTTTAAGAAAATCTGTGCTAAACTTTAATAGTTTTGTGCT a909_all.seq

      TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC Majority
               2310      2320      2330      2340      2350
298   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC 2603_all.seq
067   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC 18rs21_all.seq
299   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC coh1_all.seq
299   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC cjb111_all.seq
301   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC nem316_all.seq
300   TAATAATAATCAGCACTTACAAAGAACAAAGGGAAAAGCGAGGAGAGAAC a909_all.seq

      TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA Majority
               2360      2370      2380      2390      2400
348   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA 2603_all.seq
117   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA 18rs21_all.seq
349   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA coh1_all.seq
349   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA cjb111_all.seq
351   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA nem316_all.seq
350   TTTTAATGAAATTATCGAAGAAGTTATTGTTTTCGGCTGCTGTTTTAACA a909_all.seq

      ATGGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT Majority
               2410      2420      2430      2440      2450
398   ATGGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT 2603_all.seq
167   ATGGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT 18rs21_all.seq
399   ATGGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT coh1_all.seq
399   ATAGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT cjb111_all.seq
401   ATGGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT nem316_all.seq
400   ATAGTGGCGGGGTCAACTGTTGAACCAGTAGCTCAGTTTGCGACTGGAAT a909_all.seq
```

FIGURE 18 F

```
     GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA Majority
              2460      2470      2480      2490      2500
2448 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA 2603_all.seq
2217 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA 18rs21_all.seq
2449 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA coh1_all.seq
2449 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA cjb111_all.seq
2451 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA nem316_all.seq
2450 GAGTATTGTAAGAGCTGCAGAAGTGTCACAAGAACGCCCAGCGAAAACAA a909_all.seq

     CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT Majority
              2510      2520      2530      2540      2550
2498 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT 2603_all.seq
2267 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT 18rs21_all.seq
2499 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT coh1_all.seq
2499 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT cjb111_all.seq
2501 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT nem316_all.seq
2500 CAGTAAATATCTATAAATTACAAGCTGATAGTTATAAATCGGAAATTACT a909_all.seq

     TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC Majority
              2560      2570      2580      2590      2600
2548 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC 2603_all.seq
2317 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC 18rs21_all.seq
2549 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC coh1_all.seq
2549 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC cjb111_all.seq
2551 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC nem316_all.seq
2550 TCTAATGGTGGTATCGAGAATAAAGACGGCGAAGTAATATCTAACTATGC a909_all.seq

     TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT Majority
              2610      2620      2630      2640      2650
2598 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT 2603_all.seq
2367 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT 18rs21_all.seq
2599 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT coh1_all.seq
2599 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT cjb111_all.seq
2601 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT nem316_all.seq
2600 TAAACTTGGTGACAATGTAAAAGGTTTGCAAGGTGTACAGTTTAAACGTT a909_all.seq

     ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA Majority
              2660      2670      2680      2690      2700
2648 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA 2603_all.seq
2417 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA 18rs21_all.seq
2649 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA coh1_all.seq
2649 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA cjb111_all.seq
2651 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA nem316_all.seq
2650 ATAAAGTCAAGACGGATATTTCTGTTGATGAATTGAAAAAATTGACAACA a909_all.seq

     GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG Majority
              2710      2720      2730      2740      2750
2698 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG 2603_all.seq
2467 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG 18rs21_all.seq
2699 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG coh1_all.seq
2699 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG cjb111_all.seq
2701 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG nem316_all.seq
2700 GTTGAAGCAGCAGATGCAAAAGTTGGAACGATTCTTGAAGAAGGTGTCAG a909_all.seq

     TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT Majority
              2760      2770      2780      2790      2800
2748 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT 2603_all.seq
2517 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT 18rs21_all.seq
2749 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT coh1_all.seq
2749 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT cjb111_all.seq
2751 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT nem316_all.seq
2750 TCTACCTCAAAAAACTAATGCTCAAGGTTTGGTCGTCGATGCTCTGGATT a909_all.seq
```

FIGURE 18 G

```
      CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT Majority
               2810      2820      2830      2840      2850
2798  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT 2603_ail.seq
2567  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT 18rs21_ail.seq
2799  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT coh1_ail.seq
2799  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT cjb111_ail.seq
2801  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT nem316_ail.seq
2800  CAAAAAGTAATGTGAGATACTTGTATGTAGAAGATTTAAAGAATTCACCT a909_ail.seq

      TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT Majority
               2860      2870      2880      2890      2900
2848  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT 2603_ail.seq
2617  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT 18rs21_ail.seq
2849  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT coh1_ail.seq
2849  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT cjb111_ail.seq
2851  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT nem316_ail.seq
2850  TCAAACATTACCAAAGCTTATGCTGTACCGTTTGTGTTGGAATTACCAGT a909_ail.seq

      TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA Majority
               2910      2920      2930      2940      2950
2898  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA 2603_ail.seq
2667  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA 18rs21_ail.seq
2899  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA coh1_ail.seq
2899  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA cjb111_ail.seq
2901  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA nem316_ail.seq
2900  TGCTAACTCTACAGGTACAGGTTTCCTTTCTGAAATTAATATTTACCCTA a909_ail.seq

      AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA Majority
               2960      2970      2980      2990      3000
948   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA 2603_ail.seq
717   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAATAATTA 18rs21_ail.seq
949   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA coh1_ail.seq
949   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA cjb111_ail.seq
951   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA nem316_ail.seq
950   AAAACGTTGTAACTGATGAACCAAAAACAGATAAAGATGTTAAAAAATTA a909_ail.seq

      GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT Majority
               3010      3020      3030      3040      3050
998   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT 2603_ail.seq
767   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT 18rs21_ail.seq
999   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT coh1_ail.seq
999   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT cjb111_ail.seq
001   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT nem316_ail.seq
000   GGTCAGGACGATGCAGGTTATACGATTGGTGAAGAATTCAAATGGTTCTT a909_ail.seq

      GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA Majority
               3060      3070      3080      3090      3100
048   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA 2603_ail.seq
817   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA 18rs21_ail.seq
049   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA coh1_ail.seq
049   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA cjb111_ail.seq
051   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA nem316_ail.seq
050   GAAATCTACAATCCCTGCCAATTTAGGTGACTATGAAAAATTTGAAATTA a909_ail.seq

      CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG Majority
               3110      3120      3130      3140      3150
098   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG 2603_ail.seq
867   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG 18rs21_ail.seq
099   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG coh1_ail.seq
099   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG cjb111_ail.seq
101   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG nem316_ail.seq
100   CTGATAAATTTGCAGATGGCTTGACTTATAAATCTGTTGGAAAAATCAAG a909_ail.seq
```

FIGURE 18 H

```
    ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC Majority
         3160      3170      3180      3190      3200
148 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC 2603_ai1.seq
917 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC 18rs21_ai1.seq
149 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC coh1_ai1.seq
149 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC cjb111_ai1.seq
151 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC nem316_ai1.seq
150 ATTGGTTCGAAAACACTGAATAGAGATGAGCACTACACTATTGATGAACC a909_ai1.seq

    AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT Majority
         3210      3220      3230      3240      3250
198 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT 2603_ai1.seq
967 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT 18rs21_ai1.seq
199 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT coh1_ai1.seq
199 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT cjb111_ai1.seq
201 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT nem316_ai1.seq
200 AACAGTTGATAACCAAAATACATTAAAAATTACGTTTAAACCAGAGAAAT a909_ai1.seq

    TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA Majority
         3260      3270      3280      3290      3300
248 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA 2603_ai1.seq
017 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA 18rs21_ai1.seq
249 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA coh1_ai1.seq
249 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA cjb111_ai1.seq
251 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA nem316_ai1.seq
250 TTAAAGAAATTGCTGAGCTACTTAAAGGAATGACCCTTGTTAAAAATCAA a909_ai1.seq

    GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA Majority
         3310      3320      3330      3340      3350
298 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA 2603_ai1.seq
067 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA 18rs21_ai1.seq
299 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA coh1_ai1.seq
299 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA cjb111_ai1.seq
301 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA nem316_ai1.seq
300 GATGCTCTTGATAAAGCTACTGCAAATACAGATGATGCGGCATTTTTGGA a909_ai1.seq

    AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA Majority
         3360      3370      3380      3390      3400
48  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA 2603_ai1.seq
17  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA 18rs21_ai1.seq
49  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA coh1_ai1.seq
49  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA cjb111_ai1.seq
51  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA nem316_ai1.seq
50  AATTCCAGTTGCATCAACTATTAATGAAAAAGCAGTTTTAGGAAAAGCAA a909_ai1.seq

    TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC Majority
         3410      3420      3430      3440      3450
98  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC 2603_ai1.seq
67  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC 18rs21_ai1.seq
99  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC coh1_ai1.seq
99  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC cjb111_ai1.seq
01  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC nem316_ai1.seq
00  TTGAAAATACTTTTGAACTTCAATATGACCATACTCCTGATAAAGCTGAC a909_ai1.seq

    AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG Majority
         3460      3470      3480      3490      3500
48  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG 2603_ai1.seq
17  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG 18rs21_ai1.seq
49  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG coh1_ai1.seq
49  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG cjb111_ai1.seq
51  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG nem316_ai1.seq
50  AATCCAAAACCATCTAATCCTCCAAGAAAACCAGAAGTTCATACTGGTGG a909_ai1.seq
```

FIGURE 18 I

```
      GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG Majority
               3510      3520      3530      3540      3550
1498  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG 2603_ai1.seq
1267  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG 18rs21_ai1.seq
1499  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG coh1_ai1.seq
1499  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG cjb111_ai1.seq
1501  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG nem316_ai1.seq
1500  GAAACGATTTGTAAAGAAAGACTCAACAGAAACACAAACACTAGGTGGTG a909_ai1.seq

      CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT Majority
               3560      3570      3580      3590      3600
1548  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT 2603_ai1.seq
1317  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT 18rs21_ai1.seq
1549  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT coh1_ai1.seq
1549  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT cjb111_ai1.seq
1551  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT nem316_ai1.seq
1550  CTGAGTTTGATTTGTTGGCTTCTGATGGGACAGCAGTAAAATGGACAGAT a909_ai1.seq

      GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT Majority
               3610      3620      3630      3640      3650
598   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT 2603_ai1.seq
367   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT 18rs21_ai1.seq
599   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT coh1_ai1.seq
599   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT cjb111_ai1.seq
601   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT nem316_ai1.seq
600   GCTCTTATTAAAGCGAATACTAATAAAAACTATATTGCTGGAGAAGCTGT a909_ai1.seq

      TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA Majority
               3660      3670      3680      3690      3700
648   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA 2603_ai1.seq
417   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA 18rs21_ai1.seq
649   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA coh1_ai1.seq
649   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA cjb111_ai1.seq
651   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA nem316_ai1.seq
650   TACTGGGCAACCAATCAAATTGAAATCACATACAGACGGTACGTTTGAGA a909_ai1.seq

      TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA Majority
               3710      3720      3730      3740      3750
698   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA 2603_ai1.seq
467   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA 18rs21_ai1.seq
699   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA coh1_ai1.seq
699   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA cjb111_ai1.seq
701   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA nem316_ai1.seq
700   TTAAAGGTTTGGCTTATGCAGTTGATGCGAATGCAGAGGGTACAGCAGTA a909_ai1.seq

      ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA Majority
               3760      3770      3780      3790      3800
748   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA 2603_ai1.seq
517   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA 18rs21_ai1.seq
749   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA coh1_ai1.seq
749   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA cjb111_ai1.seq
751   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA nem316_ai1.seq
750   ACTTACAAATTAAAAGAAACAAAAGCACCAGAAGGTTATGTAATCCCTGA a909_ai1.seq

      TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA Majority
               3810      3820      3830      3840      3850
798   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA 2603_ai1.seq
567   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA 18rs21_ai1.seq
799   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA coh1_ai1.seq
799   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA cjb111_ai1.seq
801   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA nem316_ai1.seq
800   TAAAGAAATCGAGTTTACAGTATCACAAACATCTTATAATACAAAACCAA a909_ai1.seq
```

FIGURE 18 J

```
     CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC Majority
               3860      3870      3880      3890      3900
3848 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC
3617 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC 2603_ail.seq
3849 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC 18rs21_ail.seq
3849 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC coh1_ail.seq
3851 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC cjb111_ail.seq
3850 CTGACATCACGGTTGATAGTGCTGATGCAACACCTGATACAATTAAAAAC nem316_ail.seq
                                                        a909_ail.seq

     AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT Majority
               3910      3920      3930      3940      3950
3898 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT
3667 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT 2603_ail.seq
3899 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT 18rs21_ail.seq
3899 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT coh1_ail.seq
3901 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT cjb111_ail.seq
3900 AACAAACGTCCTTCAATCCCTAATACTGGTGGTATTGGTACGGCTATCTT nem316_ail.seq
                                                        a909_ail.seq

     TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC Majority
               3960      3970      3980      3990      4000
3948 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC
3717 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC 2603_ail.seq
3949 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC 18rs21_ail.seq
3949 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC coh1_ail.seq
3951 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC cjb111_ail.seq
3950 TGTCGCTATCGGTGCTGCGGTGATGGCTTTTGCTGTTAAGGGGATGAAGC nem316_ail.seq
                                                        a909_ail.seq

     GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT Majority
               4010      4020      4030      4040      4050
998  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT
767  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT 2603_ail.seq
999  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT 18rs21_ail.seq
999  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT coh1_ail.seq
001  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT cjb111_ail.seq
000  GTCGTACAAAAGATAACTAAATAAAAGGCTACTTCTTAAGTAACCATGTT nem316_ail.seq
                                                        a909_ail.seq

     TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA Majority
               4060      4070      4080      4090      4100
048  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA
817  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA 2603_ail.seq
049  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA 18rs21_ail.seq
049  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA coh1_ail.seq
051  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA cjb111_ail.seq
050  TAAGAAAAGAGAAATAGCCTTATTTCTCTTTTTGTCGTTTTTAAAATAAA nem316_ail.seq
                                                        a909_ail.seq

     GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG Majority
               4110      4120      4130      4140      4150
098  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG
867  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG 2603_ail.seq
099  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG 18rs21_ail.seq
099  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG coh1_ail.seq
101  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG cjb111_ail.seq
100  GGAACATCATGAAACAAACATTAAAACTTATGTTTTCTTTTCTGTTGATG nem316_ail.seq
                                                        a909_ail.seq

     TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA Majority
               4160      4170      4180      4190      4200
148  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA
917  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA 2603_ail.seq
149  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA 18rs21_ail.seq
149  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA coh1_ail.seq
151  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA cjb111_ail.seq
150  TTAGGGACTATGTTTGGAATTAGCCAAACTGTTTTAGCGCAAGAAACTCA nem316_ail.seq
                                                        a909_ail.seq
```

FIGURE 18 K

```
     TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC Majority
              4210      4220      4230      4240      4250
4198 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC 2603_all.seq
3967 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC 18rs21_all.seq
4199 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC coh1_all.seq
4199 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC cjb111_all.seq
4201 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC nem316_all.seq
4200 TCAGTTGACGATTGTTCATCTTGAAGCAAGGCATATTGATCGTCCAAATC a909_all.seq

     CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC Majority
              4260      4270      4280      4290      4300
4248 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC 2603_all.seq
4017 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC 18rs21_all.seq
4249 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC coh1_all.seq
4249 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC cjb111_all.seq
4251 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC nem316_all.seq
4250 CACAGTTGGAGATTGCCCCTAAAGAAGGGACTCCAATTGAAGGAGTACTC a909_all.seq

     TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA Majority
              4310      4320      4330      4340      4350
4298 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA 2603_all.seq
4067 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA 18rs21_all.seq
4299 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA coh1_all.seq
4299 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA cjb111_all.seq
4301 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA nem316_all.seq
4300 TATCAGTTGTACCAATTAAAATCAACTGAAGATGGCGATTTGTTGGCACA a909_all.seq

     TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT Majority
              4360      4370      4380      4390      4400
4348 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT 2603_all.seq
4117 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT 18rs21_all.seq
4349 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT coh1_all.seq
4349 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT cjb111_all.seq
4351 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT nem316_all.seq
4350 TTGGAATTCCCTAACTATCACAGAATTGAAAAAACAGGCGCAGCAGGTTT a909_all.seq

     TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA Majority
              4410      4420      4430      4440      4450
4398 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA 2603_all.seq
4167 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA 18rs21_all.seq
4399 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA coh1_all.seq
4399 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA cjb111_all.seq
4401 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA nem316_all.seq
4400 TTGAAGCCACTACTAATCAACAAGGAAAGGCTACATTTAACCAACTACCA a909_all.seq

     GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA Majority
              4460      4470      4480      4490      4500
4448 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA 2603_all.seq
4217 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA 18rs21_all.seq
4449 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA coh1_all.seq
4449 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA cjb111_all.seq
4451 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA nem316_all.seq
4450 GATGGAATTTATTATGGTCTGGCCGGTTAAAGCCGGTGAAAAAATCGTAA a909_all.seq

     TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA Majority
              4510      4520      4530      4540      4550
4498 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA 2603_all.seq
4267 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA 18rs21_all.seq
4499 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA coh1_all.seq
4499 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA cjb111_all.seq
4501 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA nem316_all.seq
4500 TGTCTCAGCTTTCTTGGTTGACTTGTCTGAGGATAAAGTGATTTATCCTA a909_all.seq
```

FIGURE 18 L

```
      AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT Majority
               4560      4570      4580      4590      4600
4548  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT 2603_ail.seq
4317  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT 18rs21_ail.seq
4549  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT cohl_ail.seq
4549  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT cjb111_ail.seq
4551  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT nem316_ail.seq
4550  AAATCATCTGGTCCACAGGTGAGTTGGACTTGCTTAAAGTTGGTGTGGAT a909_ail.seq

      GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA Majority
               4610      4620      4630      4640      4650
4598  GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA 2603_ail.seq
4367  GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA 18rs21_ail.seq
4599  GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA cohl_ail.seq
4599  GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA cjb111_ail.seq
4601  GGTGATACCAAAAAACCACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA nem316_ail.seq
4600  GGTGATACCAAAAAAACACTAGCAGGCGTTGTCTTTGAACTTTATGAAAA a909_ail.seq

      GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA Majority
               4660      4670      4680      4690      4700
4648  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA 2603_ail.seq
4417  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA 18rs21_ail.seq
4649  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA cohl_ail.seq
4649  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA cjb111_ail.seq
4651  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA nem316_ail.seq
4650  GAATGGTAGGACTCCTATTCGTGTGAAAAATGGGGTGCATTCTCAAGATA a909_ail.seq

      TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT Majority
               4710      4720      4730      4740      4750
4698  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT 2603_ail.seq
4467  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT 18rs21_ail.seq
4699  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT cohl_ail.seq
4699  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT cjb111_ail.seq
4701  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT nem316_ail.seq
4700  TTGACGCTGCAAAACATTTAGAAACAGATTCATCAGGGCATATCAGAATT a909_ail.seq

      TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC Majority
               4760      4770      4780      4790      4800
4748  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC 2603_ail.seq
4517  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC 18rs21_ail.seq
4749  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC cohl_ail.seq
4749  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC cjb111_ail.seq
4751  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC nem316_ail.seq
4750  TCCGGGCTCATCCATGGGGACTATGTCTTAAAAGAAATCGAGACACAGTC a909_ail.seq

      AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA Majority
               4810      4820      4830      4840      4850
4798  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA 2603_ail.seq
4567  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA 18rs21_ail.seq
4799  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA cohl_ail.seq
4799  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA cjb111_ail.seq
4801  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA nem316_ail.seq
4800  AGGATATCAGATCGGACAGGCAGAGACTGCTGTGACTATTGAAAAATCAA a909_ail.seq

      AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG Majority
               4860      4870      4880      4890      4900
4848  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG 2603_ail.seq
4617  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG 18rs21_ail.seq
4849  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG cohl_ail.seq
4849  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG cjb111_ail.seq
4851  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG nem316_ail.seq
4850  AAACAGTAACAGTAACGATTGAAAATAAAAAAGTTCCGACACCTAAAGTG a909_ail.seq
```

FIGURE 18 M

```
      CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC Majority
               4910      4920      4930      4940      4950
4898  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC 2603_ai1.seq
4667  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC 18rs21_ai1.seq
4899  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC coh1_ai1.seq
4899  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC cjb111_ai1.seq
4901  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC nem316_ai1.seq
4900  CCATCTCGAGGAGGTCTTATTCCCAAAACAGGTGAGCAACAGGCAATGGC a909_ai1.seq

      ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT Majority
               4960      4970      4980      4990      5000
4948  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT 2603_ai1.seq
4717  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT 18rs21_ai1.seq
4949  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT coh1_ai1.seq
4949  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT cjb111_ai1.seq
4951  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT nem316_ai1.seq
4950  ACTTGTAATTATTGGTGGTATTTTAATTGCTTTAGCCTTACGATTACTAT a909_ai1.seq

      CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA Majority
               5010      5020      5030      5040      5050
4998  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA 2603_ai1.seq
4767  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA 18rs21_ai1.seq
4999  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA coh1_ai1.seq
4999  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA cjb111_ai1.seq
5001  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA nem316_ai1.seq
5000  CAAAACATCGGAAACATCAAAATAAGGATTAGCATGGGACAAAAATCAAA a909_ai1.seq

      AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT Majority
               5060      5070      5080      5090      5100
5048  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT 2603_ai1.seq
4817  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT 18rs21_ai1.seq
5049  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT coh1_ai1.seq
5049  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT cjb111_ai1.seq
5051  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT nem316_ai1.seq
5050  AATATCTCTAGCTACGAATATTCGTATATGGATTTTTCGTTTAATTTTCT a909_ai1.seq

      TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC Majority
               5110      5120      5130      5140      5150
5098  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC 2603_ai1.seq
4867  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC 18rs21_ai1.seq
5099  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC coh1_ai1.seq
5099  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC cjb111_ai1.seq
5101  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC nem316_ai1.seq
5100  TAGCGGGTTTCCTTGTTTTGGCATTTCCCATCGTTAGTCAGGTCATGTAC a909_ai1.seq

      TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA Majority
               5160      5170      5180      5190      5200
148   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA 2603_ai1.seq
917   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA 18rs21_ai1.seq
149   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA coh1_ai1.seq
149   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA cjb111_ai1.seq
151   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA nem316_ai1.seq
150   TTTCAAGCCTCTCACGCCAATATTAATGCTTTTAAAGAAGCTGTTACCAA a909_ai1.seq

      GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA Majority
               5210      5220      5230      5240      5250
198   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA 2603_ai1.seq
967   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA 18rs21_ai1.seq
199   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA coh1_ai1.seq
199   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA cjb111_ai1.seq
201   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA nem316_ai1.seq
200   GATTGACCGGGTGGAGATTAATCGGCGTTTAGAACTTGCTTATGCTTATA a909_ai1.seq
```

FIGURE 18 N

```
      A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A Majority
                 5260                5270                5280                5290                5300
5248  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A 2603_all.seq
5017  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A 18rs21_all.seq
5249  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A coh1_all.seq
5249  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A cjb111_all.seq
5251  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T T C A G C G C T T A A A nem316_all.seq
5250  A C G C C A G T A T A G C A G G T G C C A A A A C T A A T G G C G A A T A T C C A G C G C T T A A A a909_all.seq

      G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G Majority
                 5310                5320                5330                5340                5350
5298  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G 2603_all.seq
5067  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G 18rs21_all.seq
5299  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G coh1_all.seq
5299  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G cjb111_all.seq
5301  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G nem316_all.seq
5300  G A C C C C T A C T C T G C T G A A C A A A A G C A G G C A G G G G T C G T T G A G T A C G C C C G a909_all.seq

      C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A Majority
                 5360                5370                5380                5390                5400
5348  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A 2603_all.seq
5117  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A 18rs21_all.seq
5349  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A coh1_all.seq
5349  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A cjb111_all.seq
5351  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A nem316_all.seq
5350  C A T G C T T G A A G T C A A A G A A C A A A T A G G T C A T G T G A T T A T T C C A A G A A T T A a909_all.seq

      A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G Majority
                 5410                5420                5430                5440                5450
398   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G 2603_all.seq
167   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G 18rs21_all.seq
399   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G coh1_all.seq
399   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G cjb111_all.seq
401   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G nem316_all.seq
400   A T C A G G A T A T C C C T A T T T A C G C T G G C T C T G C T G A A G A A A A T C T T C A G A G G a909_all.seq

      G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C Majority
                 5460                5470                5480                5490                5500
448   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C 2603_all.seq
217   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C 18rs21_all.seq
449   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C coh1_all.seq
449   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C cjb111_all.seq
451   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C nem316_all.seq
450   G G C G T T G G A C A T T T A G A G G G G A C C A G T C T T C C A G T C G G T G G T G A G T C A A C a909_all.seq

      T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A Majority
                 5510                5520                5530                5540                5550
498   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A 2603_all.seq
267   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A 18rs21_all.seq
499   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A coh1_all.seq
499   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A cjb111_all.seq
501   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A nem316_all.seq
500   T C A T G C C G T T C T A A C T G C C C A T C G A G G G C T A C C A A C G G C C A A G C T A T T T A a909_all.seq

      C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C Majority
                 5560                5570                5580                5590                5600
548   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C 2603_all.seq
317   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C 18rs21_all.seq
549   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C coh1_all.seq
549   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C cjb111_all.seq
551   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C nem316_all.seq
550   C C A A T T T A G A C A A G G T A A C A G T A G G T G A C C G T T T T T A C A T T G A A C A C A T C a909_all.seq
```

FIGURE 18 O

```
      GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA Majority
               5610      5620      5630      5640      5650
598   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA 2603_ail.seq
367   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA 18rs21_ail.seq
599   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA coh1_ail.seq
599   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA cjb111_ail.seq
601   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA nem316_ail.seq
600   GCCGGAAAGATTGCTTATCAGGTAGACCAAATCAAAGTTATCGCCCCTGA a909_ail.seq

      TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT Majority
               5660      5670      5680      5690      5700
648   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT 2603_ail.seq
417   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT 18rs21_ail.seq
649   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT coh1_ail.seq
649   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT cjb111_ail.seq
651   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT nem316_ail.seq
650   TCAGTTAGAGGATTTGTACGTGATTCAAGGAGAAGATCACGTCACCCTAT a909_ail.seq

      TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC Majority
               5710      5720      5730      5740      5750
698   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC 2603_ail.seq
467   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC 18rs21_ail.seq
699   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC coh1_ail.seq
699   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC cjb111_ail.seq
701   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC nem316_ail.seq
700   TAACTTGCACACCTTATATGATAAATAGTCATCGCCTCCTCGTTCGAGGC a909_ail.seq

      AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT Majority
               5760      5770      5780      5790      5800
748   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT 2603_ail.seq
517   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT 18rs21_ail.seq
749   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT coh1_ail.seq
749   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT cjb111_ail.seq
751   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT nem316_ail.seq
750   AAGCGAATTCCTTATGTGGAAAAAACAGTGCAGAAAGATTCAAAGACCTT a909_ail.seq

      CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA Majority
               5810      5820      5830      5840      5850
798   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA 2603_ail.seq
567   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA 18rs21_ail.seq
799   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA coh1_ail.seq
799   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA cjb111_ail.seq
801   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA nem316_ail.seq
800   CAGGCAACAACAATACCTAACCTATGCTATGTGGGTAGTCGTTGGACTTA a909_ail.seq

      TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG Majority
               5860      5870      5880      5890      5900
848   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG 2603_ail.seq
617   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG 18rs21_ail.seq
849   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG coh1_ail.seq
849   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG cjb111_ail.seq
851   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG nem316_ail.seq
850   TCTTGCTGTCGCTTCTCATTTGGTTTAAAAAGACGAAACAGAAAAAGCGG a909_ail.seq

      AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA Majority
               5910      5920      5930      5940      5950
898   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA 2603_ail.seq
667   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA 18rs21_ail.seq
899   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA coh1_ail.seq
899   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA cjb111_ail.seq
901   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA nem316_ail.seq
900   AGAAAGAATGAAAAAGCGGCTAGTCAAAATAGTCACAATAATTCGAAATA a909_ail.seq
```

FIGURE 18 P

```
      ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT Majority
               5960      5970      5980      5990      6000
5948  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT 2603_ail.seq
5717  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT 18rs21_ail.seq
5949  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT coh1_ail.seq
5949  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT cjb111_ail.seq
5951  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT nem316_ail.seq
5950  ATAAAATCAGAACCCTCATTTTTGTGATGGCAAGTCTGATTCTCTTATTT a909_ail.seq

      CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA Majority
               6010      6020      6030      6040      6050
5998  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA 2603_ail.seq
5767  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA 18rs21_ail.seq
5999  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA coh1_ail.seq
5999  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA cjb111_ail.seq
6001  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA nem316_ail.seq
6000  CCGATTGTGAGCCAGGTAAGTTACTACCTTGCTTCGCATCAAAATATTAA a909_ail.seq

      TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC Majority
               6060      6070      6080      6090      6100
6048  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC 2603_ail.seq
5817  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC 18rs21_ail.seq
6049  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC coh1_ail.seq
6049  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC cjb111_ail.seq
6051  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC nem316_ail.seq
6050  TCAATTTAAGCGGGAAGTCGCTAAGATTGATACTAATACGGTTGAACGAC a909_ail.seq

      GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG Majority
               6110      6120      6130      6140      6150
6098  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG 2603_ail.seq
5867  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG 18rs21_ail.seq
6099  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG coh1_ail.seq
6099  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG cjb111_ail.seq
6101  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG nem316_ail.seq
6100  GCATCGCTTTAGCTAATGCTTACAATGAGACGTTATCAAGGAATCCCTTG a909_ail.seq

      CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC Majority
               6160      6170      6180      6190      6200
148   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC 2603_ail.seq
917   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC 18rs21_ail.seq
149   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC coh1_ail.seq
149   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC cjb111_ail.seq
151   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC nem316_ail.seq
150   CTTATAGACCCTTTTACCAGTAAGCAAAAAGAAGGTTTGAGAGAGTATGC a909_ail.seq

      TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA Majority
               6210      6220      6230      6240      6250
198   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA 2603_ail.seq
967   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA 18rs21_ail.seq
199   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA coh1_ail.seq
199   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA cjb111_ail.seq
201   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA nem316_ail.seq
200   TCGTATGCTTGAAGTTCATGAGCAAATAGGTCATGTGGCAATCCCAAGTA a909_ail.seq

      TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG Majority
               6260      6270      6280      6290      6300
248   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG 2603_ail.seq
017   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG 18rs21_ail.seq
249   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG coh1_ail.seq
249   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG cjb111_ail.seq
251   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG nem316_ail.seq
250   TTGCGGTTGATATTCCAATTTATGCTGGAACATCCGAAACTGTGCTTCAG a909_ail.seq
```

FIGURE 18 Q

	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	Majority
	6310 6320 6330 6340 6350	
6298	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	2603_ail.seq
6067	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	18rs21_ail.seq
6299	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	coh1_ail.seq
6299	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	cjb111_ail.seq
6301	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	nem316_ail.seq
6300	AAAGGTAGTGGGCATTTGGAGGGAACCAGTCTTCCAGTGGGAGGTTTGTC	a909_ail.seq

	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	Majority
	6360 6370 6380 6390 6400	
6348	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	2603_ail.seq
6117	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	18rs21_ail.seq
6349	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	coh1_ail.seq
6349	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	cjb111_ail.seq
6351	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	nem316_ail.seq
6350	AACCCATTCAGTACTAACTGCCCACCGTGGCTTGCCAACAGCTAGGCTAT	a909_ail.seq

	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	Majority
	6410 6420 6430 6440 6450	
6398	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	2603_ail.seq
6167	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	18rs21_ail.seq
6399	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	coh1_ail.seq
6399	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	cjb111_ail.seq
6401	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	nem316_ail.seq
6400	TTACCGACTTAAATAAAGTTAAAAAAGGCCAGATTTTCTATGTGACGAAC	a909_ail.seq

	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	Majority
	6460 6470 6480 6490 6500	
6448	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	2603_ail.seq
6217	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	18rs21_ail.seq
6449	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	coh1_ail.seq
6449	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	cjb111_ail.seq
6451	ATCAAAGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	nem316_ail.seq
6450	ATCAAGGAAACACTTGCCTACAAAGTCGTGTCTATCAAAGTTGTGGATCC	a909_ail.seq

	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	Majority
	6510 6520 6530 6540 6550	
6498	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	2603_ail.seq
6267	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	18rs21_ail.seq
6499	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	coh1_ail.seq
6499	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	cjb111_ail.seq
6501	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	nem316_ail.seq
6500	AACAGCTTTAAGTGAGGTTAAGATTGTCAATGGTAAGGATTATATAACCT	a909_ail.seq

	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	Majority
	6560 6570 6580 6590 6600	
6548	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	2603_ail.seq
6317	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	18rs21_ail.seq
6549	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	coh1_ail.seq
6549	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	cjb111_ail.seq
6551	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	nem316_ail.seq
6550	TGCTGACTTGCACACCTTACATGATCAATAGTCATCGTCTCTTGGTAAAA	a909_ail.seq

	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	Majority
	6610 6620 6630 6640 6650	
6598	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	2603_ail.seq
6367	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	18rs21_ail.seq
6599	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	coh1_ail.seq
6599	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	cjb111_ail.seq
6601	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	nem316_ail.seq
6600	GGAGAGCCGTATTCCTTATGATTCTACCGAGGCGGAAAAGCACAAAGAACA	a909_ail.seq

FIGURE 18 R

```
      AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT Majority
               6660      6670      6680      6690      6700
6648  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT 2603_all.seq
6417  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT 18rs21_all.seq
6649  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT coh1_all.seq
6649  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT cjb111_all.seq
6651  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT nem316_all.seq
6650  AACCGTACAAGATTATCGTTTGTCACTAGTGTTGAAGATACTACTAGTAT a909_all.seq

      TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT Majority
               6710      6720      6730      6740      6750
6698  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT 2603_all.seq
6467  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT 18rs21_all.seq
6699  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT coh1_all.seq
6699  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT cjb111_all.seq
6701  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT nem316_all.seq
6700  TATTAATTGGACTCTTCATCGTGATAATGATGAGAAGATGGATGCAACAT a909_all.seq

      CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT Majority
               6760      6770      6780      6790      6800
6748  CGTCAATAACGATGTTGTGAATGGCTTACTT----ATCAAATAGGTGACT 2603_all.seq
6517  CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT 18rs21_all.seq
6749  CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT coh1_all.seq
6749  CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT cjb111_all.seq
6751  CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT nem316_all.seq
6750  CGTCAATAACGATGTTGTGAATGGCTTACTTACTTATCAAATAGGTGACT a909_all.seq

      AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA Majority
               6810      6820      6830      6840      6850
6794  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA 2603_all.seq
6567  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA 18rs21_all.seq
6799  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA coh1_all.seq
6799  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA cjb111_all.seq
6801  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA nem316_all.seq
6800  AATGATGATTGTGAATAATGGTTATCTAGAAGGGAGAAAAATGAAAAAGA a909_all.seq

      GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA Majority
               6860      6870      6880      6890      6900
6844  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA 2603_all.seq
6617  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA 18rs21_all.seq
6849  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA coh1_all.seq
6849  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA cjb111_all.seq
6851  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA nem316_all.seq
6850  GACAAAAAATATGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAA a909_all.seq

      ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC Majority
               6910      6920      6930      6940      6950
6894  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC 2603_all.seq
6667  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC 18rs21_all.seq
6899  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC coh1_all.seq
6899  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC cjb111_all.seq
6901  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC nem316_all.seq
6900  ATTCCATTTGGTATATTGGTACAAGGTGAAACCCAAGATACCAATCAAGC a909_all.seq

      ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG Majority
               6960      6970      6980      6990      7000
6944  ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG 2603_all.seq
6717  ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG 18rs21_all.seq
6949  ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG coh1_all.seq
6949  ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG cjb111_all.seq
6951  ACTTGGAAAAGTAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAG nem316_all.seq
6950  ACTTGGAAAAGTAATTGTTAAAAAAACGGGGGACAATGCTACACCATTAG a909_all.seq
```

FIGURE 18 S

	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	Majority
	7010 7020 7030 7040 7050	
6994	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	2603_ai1.seq
6767	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	18rs21_ai1.seq
6999	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	coh1_ai1.seq
6999	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	cjb111_ai1.seq
7001	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	nem316_ai1.seq
7000	G C A A A G C G A C T T T T G T G T T A A A A A A T G A C A A T G A T A A G T C A G A A A C A A G T	a909_ai1.seq
	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	Majority
	7060 7070 7080 7090 7100	
7044	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	2603_ai1.seq
6817	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	18rs21_ai1.seq
7049	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	coh1_ai1.seq
7049	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	cjb111_ai1.seq
7051	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	nem316_ai1.seq
7050	C A C G A A A C G G T A G A G G G T T C T G G A G A A G C A A C C T T T G A A A A C A T A A A A C C	a909_ai1.seq
	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	Majority
	7110 7120 7130 7140 7150	
7094	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	2603_ai1.seq
6867	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	18rs21_ai1.seq
7099	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	coh1_ai1.seq
7099	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	cjb111_ai1.seq
7101	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	nem316_ai1.seq
7100	T G G A G A C T A C A C A T T A A G A G A A G A A A C A G C A C C A A T T G G T T A T A A A A A A A	a909_ai1.seq
	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	Majority
	7160 7170 7180 7190 7200	
7144	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	2603_ai1.seq
6917	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	18rs21_ai1.seq
7149	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	coh1_ai1.seq
7149	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	cjb111_ai1.seq
7151	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	nem316_ai1.seq
7150	C T G A T A A A A C C T G G A A A G T T A A A G T T G C A G A T A A C G G A G C A A C A A T A A T C	a909_ai1.seq
	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	Majority
	7210 7220 7230 7240 7250	
7194	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	2603_ai1.seq
6967	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	18rs21_ai1.seq
7199	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	coh1_ai1.seq
7199	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	cjb111_ai1.seq
7201	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	nem316_ai1.seq
7200	G A G G G T A T G G A T G C A G A T A A A G C A G A G A A A C G A A A A G A A G T T T T G A A T G C	a909_ai1.seq
	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	Majority
	7260 7270 7280 7290 7300	
7244	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	2603_ai1.seq
7017	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	18rs21_ai1.seq
7249	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	coh1_ai1.seq
7249	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	cjb111_ai1.seq
7251	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	nem316_ai1.seq
7250	C C A A T A T C C A A A A T C A G C T A T T T A T G A G G A T A C A A A A G A A A A T T A C C C A T	a909_ai1.seq
	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	Majority
	7310 7320 7330 7340 7350	
7294	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	2603_ai1.seq
7067	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	18rs21_ai1.seq
7299	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	coh1_ai1.seq
7299	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	cjb111_ai1.seq
7301	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	nem316_ai1.seq
7300	T A G T T A A T G T A G A G G G T T C C A A A G T T G G T G A A C A A T A C A A A G C A T T G A A T	a909_ai1.seq

FIGURE 18 T

	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	Majority
	7360 7370 7380 7390 7400	
7344	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	2603_ai1.seq
7117	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	18rs21_ai1.seq
7349	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	coh1_ai1.seq
7349	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	cjb111_ai1.seq
7351	CCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATC	nem316_ai1.seq
7350	CCAATAAATGGAAAAGATGGTCGAAGACAGATTGCTGAAGGTTGGTTATC	a909_ai1.seq
	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	Majority
	7410 7420 7430 7440 7450	
7394	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	2603_ai1.seq
7167	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	18rs21_ai1.seq
7399	AAAAAAAAATACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	coh1_ai1.seq
7399	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	cjb111_ai1.seq
7401	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	nem316_ai1.seq
7400	AAAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTG	a909_ai1.seq
	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	Majority
	7460 7470 7480 7490 7500	
7444	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	2603_ai1.seq
7217	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	18rs21_ai1.seq
7449	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	coh1_ai1.seq
7449	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	cjb111_ai1.seq
7451	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	nem316_ai1.seq
7450	AATTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAA	a909_ai1.seq
	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	Majority
	7510 7520 7530 7540 7550	
494	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	2603_ai1.seq
267	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	18rs21_ai1.seq
499	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	coh1_ai1.seq
499	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	cjb111_ai1.seq
501	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	nem316_ai1.seq
500	CCACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGA	a909_ai1.seq
	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	Majority
	7560 7570 7580 7590 7600	
544	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	2603_ai1.seq
317	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	18rs21_ai1.seq
549	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	coh1_ai1.seq
549	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	cjb111_ai1.seq
551	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	nem316_ai1.seq
550	AAGAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAA	a909_ai1.seq
	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	Majority
	7610 7620 7630 7640 7650	
594	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	2603_ai1.seq
367	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	18rs21_ai1.seq
599	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	coh1_ai1.seq
599	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	cjb111_ai1.seq
601	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	nem316_ai1.seq
600	AGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTG	a909_ai1.seq
	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	Majority
	7660 7670 7680 7690 7700	
644	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	2603_ai1.seq
417	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	18rs21_ai1.seq
649	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	coh1_ai1.seq
649	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	cjb111_ai1.seq
651	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	nem316_ai1.seq
650	ACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGG	a909_ai1.seq

FIGURE 18 U

```
      AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT Majority
               7710      7720      7730      7740      7750
7694  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT 2603_all.seq
7467  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT 18rs21_all.seq
7699  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT cohl_all.seq
7699  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT cjb111_all.seq
7701  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT nem316_all.seq
7700  AGTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATT a909_all.seq

      ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT Majority
               7760      7770      7780      7790      7800
7744  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT 2603_all.seq
7517  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT 18rs21_all.seq
7749  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT cohl_all.seq
7749  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT cjb111_all.seq
7751  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT nem316_all.seq
7750  ATCATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAAT a909_all.seq

      TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA Majority
               7810      7820      7830      7840      7850
7794  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA 2603_all.seq
7567  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA 18rs21_all.seq
7799  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA cohl_all.seq
7799  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA cjb111_all.seq
7801  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA nem316_all.seq
7800  TTAACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAA a909_all.seq

      GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA Majority
               7860      7870      7880      7890      7900
7844  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA 2603_all.seq
7617  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA 18rs21_all.seq
7849  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA cohl_all.seq
7849  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA cjb111_all.seq
7851  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA nem316_all.seq
7850  GGAAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGA a909_all.seq

      CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA Majority
               7910      7920      7930      7940      7950
7894  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA 2603_all.seq
7667  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA 18rs21_all.seq
7899  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA cohl_all.seq
7899  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA cjb111_all.seq
7901  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA nem316_all.seq
7900  CATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAA a909_all.seq

      AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC Majority
               7960      7970      7980      7990      8000
7944  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC 2603_all.seq
7717  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC 18rs21_all.seq
7949  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC cohl_all.seq
7949  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC cjb111_all.seq
7951  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC nem316_all.seq
7950  AGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCC a909_all.seq

      TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC Majority
               8010      8020      8030      8040      8050
7994  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC 2603_all.seq
7767  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC 18rs21_all.seq
7999  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC cohl_all.seq
7999  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC cjb111_all.seq
8001  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC nem316_all.seq
8000  TACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACC a909_all.seq
```

FIGURE 18 V

	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	Majority
	8060 8070 8080 8090 8100	
3044	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	2603_all.seq
7817	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	18rs21_all.seq
3049	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	coh1_all.seq
3049	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	cjb111_all.seq
3051	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	nem316_all.seq
3050	AAAACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATT	a909_all.seq
	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	Majority
	8110 8120 8130 8140 8150	
3094	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	2603_all.seq
7867	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	18rs21_all.seq
3099	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	coh1_all.seq
3099	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	cjb111_all.seq
3101	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	nem316_all.seq
3100	CTCCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGG	a909_all.seq
	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	Majority
	8160 8170 8180 8190 8200	
3144	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	2603_all.seq
7917	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	18rs21_all.seq
3149	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	coh1_all.seq
3149	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	cjb111_all.seq
3151	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	nem316_all.seq
3150	AGATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTG	a909_all.seq
	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	Majority
	8210 8220 8230 8240 8250	
194	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	2603_all.seq
967	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	18rs21_all.seq
199	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	coh1_all.seq
199	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	cjb111_all.seq
201	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	nem316_all.seq
200	GAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTA	a909_all.seq
	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	Majority
	8260 8270 8280 8290 8300	
244	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	2603_all.seq
017	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	18rs21_all.seq
249	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	coh1_all.seq
249	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	cjb111_all.seq
251	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	nem316_all.seq
250	ATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTG	a909_all.seq
	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	Majority
	8310 8320 8330 8340 8350	
294	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	2603_all.seq
067	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	18rs21_all.seq
299	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	coh1_all.seq
299	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	cjb111_all.seq
301	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	nem316_all.seq
300	GAGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTT	a909_all.seq
	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	Majority
	8360 8370 8380 8390 8400	
344	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	2603_all.seq
117	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	18rs21_all.seq
349	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	coh1_all.seq
349	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	cjb111_all.seq
351	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	nem316_all.seq
350	CTGCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTT	a909_all.seq

FIGURE 18 W

	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	Majority
	8410 8420 8430 8440 8450	
8394	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	2603_all.seq
8167	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	18rs21_all.seq
8399	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	coh1_all.seq
8399	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	cjb111_all.seq
8401	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	nem316_all.seq
8400	AATGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGG	a909_all.seq

	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	Majority
	8460 8470 8480 8490 8500	
8444	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	2603_all.seq
8217	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	18rs21_all.seq
8449	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	coh1_all.seq
8449	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	cjb111_all.seq
8451	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	nem316_all.seq
8450	TGTAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGC	a909_all.seq

	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	Majority
	8510 8520 8530 8540 8550	
8494	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	2603_all.seq
8267	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	18rs21_all.seq
8499	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	coh1_all.seq
8499	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	cjb111_all.seq
3501	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	nem316_all.seq
3500	AATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAAT	a909_all.seq

	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	Majority
	8560 8570 8580 8590 8600	
3544	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	2603_all.seq
3317	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	18rs21_all.seq
3549	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	coh1_all.seq
3549	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	cjb111_all.seq
3551	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	nem316_all.seq
3550	AAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAA	a909_all.seq

	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	Majority
	8610 8620 8630 8640 8650	
1594	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	2603_all.seq
1367	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	18rs21_all.seq
1599	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	coh1_all.seq
1599	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	cjb111_all.seq
1601	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	nem316_all.seq
1600	ACATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTG	a909_all.seq

	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	Majority
	8660 8670 8680 8690 8700	
644	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	2603_all.seq
417	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	18rs21_all.seq
649	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	coh1_all.seq
649	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	cjb111_all.seq
651	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	nem316_all.seq
650	AATTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTG	a909_all.seq

	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	Majority
	8710 8720 8730 8740 8750	
694	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	2603_all.seq
467	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	18rs21_all.seq
699	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	coh1_all.seq
699	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	cjb111_all.seq
701	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	nem316_all.seq
700	GTTGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACC	a909_all.seq

FIGURE 18 X

```
     AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA Majority
              8760      8770      8780      8790      8800
8744 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA 2603_all.seq
8517 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA 18rs21_all.seq
8749 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA coh1_all.seq
8749 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA cjb111_all.seq
8751 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA nem316_all.seq
8750 AAACAGTGATGGCGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGA a909_all.seq

     CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA Majority
              8810      8820      8830      8840      8850
8794 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA 2603_all.seq
8567 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA 18rs21_all.seq
8799 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA coh1_all.seq
8799 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA cjb111_all.seq
8801 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA nem316_all.seq
8800 CATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAA a909_all.seq

     GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA Majority
              8860      8870      8880      8890      8900
8844 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA 2603_all.seq
8617 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA 18rs21_all.seq
8849 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA coh1_all.seq
8849 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA cjb111_all.seq
8851 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA nem316_all.seq
8850 GTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAA a909_all.seq

     ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG Majority
              8910      8920      8930      8940      8950
8894 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG 2603_all.seq
8667 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG 18rs21_all.seq
8899 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG coh1_all.seq
8899 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG cjb111_all.seq
8901 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG nem316_all.seq
8900 ATTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAG a909_all.seq

     AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT Majority
              8960      8970      8980      8990      9000
8944 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT 2603_all.seq
8717 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT 18rs21_all.seq
8949 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT coh1_all.seq
8949 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT cjb111_all.seq
8951 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT nem316_all.seq
8950 AACCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGT a909_all.seq

     GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA Majority
              9010      9020      9030      9040      9050
8994 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA 2603_all.seq
8767 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA 18rs21_all.seq
8999 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA coh1_all.seq
8999 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA cjb111_all.seq
9001 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA nem316_all.seq
9000 GAGTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGA a909_all.seq

     ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA Majority
              9060      9070      9080      9090      9100
9044 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA 2603_all.seq
8817 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA 18rs21_all.seq
9049 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA coh1_all.seq
9049 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA cjb111_all.seq
9051 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA nem316_all.seq
9050 ATTTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGAGCTA a909_all.seq
```

FIGURE 18 Y

```
     AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT Majority
          9110      9120      9130      9140      9150
9094 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT 2603_all.seq
8867 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT 18rs21_all.seq
9099 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT coh1_all.seq
9099 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT cjb111_all.seq
9101 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT nem316_all.seq
9100 AGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTT a909_all.seq

     CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA Majority
          9160      9170      9180      9190      9200
9144 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA 2603_all.seq
8917 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA 18rs21_all.seq
9149 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA coh1_all.seq
9149 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA cjb111_all.seq
9151 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA nem316_all.seq
9150 CCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAA a909_all.seq

     AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG Majority
          9210      9220      9230      9240      9250
9194 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG 2603_all.seq
8967 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG 18rs21_all.seq
9199 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG coh1_all.seq
9199 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG cjb111_all.seq
9201 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG nem316_all.seq
9200 AGCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATG a909_all.seq

     GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT Majority
          9260      9270      9280      9290      9300
9244 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT 2603_all.seq
9017 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT 18rs21_all.seq
9249 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT coh1_all.seq
9249 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT cjb111_all.seq
9251 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT nem316_all.seq
9250 GCTATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAAT a909_all.seq

     GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT Majority
          9310      9320      9330      9340      9350
294  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT 2603_all.seq
067  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT 18rs21_all.seq
299  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT coh1_all.seq
299  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT cjb111_all.seq
301  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT nem316_all.seq
300  GGAGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAAT a909_all.seq

     CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC Majority
          9360      9370      9380      9390      9400
344  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC 2603_all.seq
117  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC 18rs21_all.seq
349  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC coh1_all.seq
349  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC cjb111_all.seq
351  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC nem316_all.seq
350  CGGGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAAC a909_all.seq

     GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT Majority
          9410      9420      9430      9440      9450
394  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT 2603_all.seq
167  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT 18rs21_all.seq
399  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT coh1_all.seq
399  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT cjb111_all.seq
401  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT nem316_all.seq
400  GCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACAATTGTCTAT a909_all.seq
```

FIGURE 18 Z

```
     ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG Majority
              9460      9470      9480      9490      9500
9444 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG 2603_all.seq
9217 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG 18rs21_all.seq
9449 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG coh1_all.seq
9449 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG cjb111_all.seq
9451 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG nem316_all.seq
9450 ATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCG a909_all.seq

     TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA Majority
              9510      9520      9530      9540      9550
9494 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA 2603_all.seq
9267 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA 18rs21_all.seq
9499 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA coh1_all.seq
9499 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA cjb111_all.seq
9501 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA nem316_all.seq
9500 TAAACAATTGTAAGGTGTCGTTGAAATTATTAAATAATAGAAAATGACTA a909_all.seq

     GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC Majority
              9560      9570      9580      9590      9600
9544 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC 2603_all.seq
9317 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC 18rs21_all.seq
9549 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC coh1_all.seq
9549 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC cjb111_all.seq
9551 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC nem316_all.seq
9550 GTTTTGGCCTTTCCCTATTGTCAGTCAGATTAGTTATTATCAAGCTTCTC a909_all.seq

     ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG Majority
              9610      9620      9630      9640      9650
 594 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG 2603_all.seq
 367 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG 18rs21_all.seq
 599 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG coh1_all.seq
 599 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG cjb111_all.seq
 601 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG nem316_all.seq
 600 ATGCCAATATTAATGCCTTTAAAAGAACAGTTACAACCATTGACCGTACG a909_all.seq

     GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC Majority
              9660      9670      9680      9690      9700
 644 GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC 2603_all.seq
 417 GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC 18rs21_all.seq
 649 GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC coh1_all.seq
 649 GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC cjb111_all.seq
 651 GAGATTCAACGACGGTTAGGTCTGTCTAAAGCCTACAATGCTAGTATTTC nem316_all.seq
 650 GAGATTCAACGACGGTTAGGTCTGGCTAAAGCCTACAATGCTAGTATTTC a909_all.seq

     TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG Majority
              9710      9720      9730      9740      9750
 694 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG 2603_all.seq
 467 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG 18rs21_all.seq
 699 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG coh1_all.seq
 699 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG cjb111_all.seq
 701 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG nem316_all.seq
 700 TGGAACAAGTAGTCAGTCGACTCAATCTGTGCTGAGAGATTCTTATTCTG a909_all.seq

     AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA Majority
              9760      9770      9780      9790      9800
 744 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA 2603_all.seq
 517 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA 18rs21_all.seq
 749 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA coh1_all.seq
 749 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA cjb111_all.seq
 751 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA nem316_all.seq
 750 AGGAGCAAAAAAGGCAGGGTGGACTGAATACGCTAGGATGTTAGAAGTCA a909_all.seq
```

FIGURE 18 AA

```
      GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA Majority
               9810      9820      9830      9840      9850
794   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA 2603_ail.seq
567   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA 18rs21_ail.seq
799   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA coh1_ail.seq
799   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA cjb111_ail.seq
801   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA nem316_ail.seq
800   GAGAGCAGGTTGACCATGTGATGATTCCAAAAATCAATCAGGATTTACCA a909_ail.seq

      ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT Majority
               9860      9870      9880      9890      9900
844   ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT 2603_ail.seq
617   ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT 18rs21_ail.seq
849   ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT coh1_ail.seq
849   ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT cjb111_ail.seq
851   ATCTACGCTGGTCCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT nem316_ail.seq
850   ATCTACGCTGGTTCAGAAGAGGACAATCTGCAACGGGGAGTTGGTCATCT a909_ail.seq

      AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA Majority
               9910      9920      9930      9940      9950
894   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA 2603_ail.seq
667   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA 18rs21_ail.seq
899   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA coh1_ail.seq
899   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA cjb111_ail.seq
901   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA nem316_ail.seq
900   AGAAGGGATAAGTTTGCCGATTGGAGGGGCTTCTACACATGCGGTCTTGA a909_ail.seq

      GCGGTCAAAGAGGTATGCCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG Majority
               9960      9970      9980      9990      10000
944   GCGGTCAAAGAGGTATGCCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG 2603_ail.seq
717   GCGGTCAAAGAGGTATGCCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG 18rs21_ail.seq
949   GCGGTCAAAGAGGTATGCCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG coh1_ail.seq
949   GCGGTCAAAGAGGTATGTCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG cjb111_ail.seq
951   GCGGTCAAAGAGGTATGCCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG nem316_ail.seq
950   GCGGTCAAAGAGGTATGTCAGCTGCTCGGTTGTTTGCGGATTTGGATAAG a909_ail.seq

      ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC Majority
              10010     10020     10030     10040     10050
994   ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC 2603_ail.seq
767   ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC 18rs21_ail.seq
999   ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC coh1_ail.seq
999   ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC cjb111_ail.seq
0001  ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC nem316_ail.seq
0000  ATGAAAAAAGGTGATTATTTTTATGTTACCAATCTGAAAGAAACCTTGGC a909_ail.seq

      TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG Majority
              10060     10070     10080     10090     10100
0044  TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG 2603_ail.seq
817   TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG 18rs21_ail.seq
0049  TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG coh1_ail.seq
0049  TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG cjb111_ail.seq
0051  TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG nem316_ail.seq
0050  TTATCAAGTGGATCGTATCATGGTGATTGAACCTAGCCAATTGGATGCCG a909_ail.seq

      TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT Majority
              10110     10120     10130     10140     10150
0094  TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT 2603_ail.seq
867   TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT 18rs21_ail.seq
0099  TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT coh1_ail.seq
0099  TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT cjb111_ail.seq
0101  TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT nem316_ail.seq
0100  TGAGCATTGAAGAGGATAAAGATTATGTTACCCTTCTGACCTGTACACCT a909_ail.seq
```

FIGURE 18 AB

```
      TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG Majority
                10160     10170     10180     10190     10200
10144 TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG 2603_all.seq
9917  TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG 18rs21_all.seq
10149 TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG coh1_all.seq
10149 TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG cjb111_all.seq
10151 TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG nem316_all.seq
10150 TATATGGGCTCTTTGTCAACTGTAATGGGTGACTTATCATTAACAACGAG a909_all.seq

      AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA Majority
                10210     10220     10230     10240     10250
10194 AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA 2603_all.seq
9967  AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA 18rs21_all.seq
10199 AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA coh1_all.seq
10199 AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA cjb111_all.seq
10201 AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA nem316_all.seq
10200 AGAGAATCAGCTTGGTTCTCTCTCTTTTTGGATGTTCAAAGCGATGAGAA a909_all.seq

      TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG Majority
                10260     10270     10280     10290     10300
10244 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG 2603_all.seq
10017 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG 18rs21_all.seq
10249 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG coh1_all.seq
10249 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG cjb111_all.seq
10251 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG nem316_all.seq
10250 TTTTGCTTTTAAAATTTTTAAAGTTGCGAAAGCCAAAGGCTTGTCGCTTG a909_all.seq

      ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA Majority
                10310     10320     10330     10340     10350
10294 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA 2603_all.seq
10067 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA 18rs21_all.seq
10299 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA coh1_all.seq
10299 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA cjb111_all.seq
10301 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA nem316_all.seq
10300 ATGTCTTTGATAAGCTTGTTGGTCGCTTCTAATTTGGCATTGGAAAGGGA a909_all.seq

      TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA Majority
                10360     10370     10380     10390     10400
10344 TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA 2603_all.seq
10117 TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA 18rs21_all.seq
10349 TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA coh1_all.seq
10349 TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA cjb111_all.seq
0351  TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA nem316_all.seq
0350  TAGCTGAGGGCGTTTGTGATGTAAGTTTTGTACTTCATGAAGGTCCTAAA a909_all.seq

      GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC Majority
                10410     10420     10430     10440     10450
0394  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC 2603_all.seq
0167  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC 18rs21_all.seq
0399  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC coh1_all.seq
0399  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC cjb111_all.seq
0401  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC nem316_all.seq
0400  GACCATTTTAAAAGCAGAATTGCCAGTTGTTACGTTATCATCTATGAAGC a909_all.seq

      TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA Majority
                10460     10470     10480     10490     10500
0444  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA 2603_all.seq
0217  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA 18rs21_all.seq
0449  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA coh1_all.seq
0449  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA cjb111_all.seq
0451  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA nem316_all.seq
0450  TAAAGAATTCAGTTGTTCGCTTTTCTTAGAAATGAAAGAGTATGAGCTGA a909_all.seq
```

FIGURE 18 AC

```
      TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA Majority
               10510     10520     10530     10540     10550
10494 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA 2603_all.seq
10267 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA 18rs21_all.seq
10499 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA coh1_all.seq
10499 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA cjb111_all.seq
10501 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA nem316_all.seq
10500 TAAAGGTCGTTATAATAGCGGAGCTCATCTGAGAAAACTTTTATACCTCA a909_all.seq

      AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC Majority
               10560     10570     10580     10590     10600
10544 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC 2603_all.seq
10317 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC 18rs21_all.seq
10549 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC coh1_all.seq
10549 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC cjb111_all.seq
10551 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC nem316_all.seq
10550 AAGTCAGTCTAGCTTTGATATCATGAAGCCATTAGGAGTTATTCCTTATC a909_all.seq

      TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA Majority
               10610     10620     10630     10640     10650
10594 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA 2603_all.seq
10367 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA 18rs21_all.seq
10599 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA coh1_all.seq
10599 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA cjb111_all.seq
10601 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA nem316_all.seq
10600 TTTTAGTGGCGCGCGATCCATATAGTGATAGATCGAGATATTTAGATCCA a909_all.seq

      AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA Majority
               10660     10670     10680     10690     10700
10644 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA 2603_all.seq
10417 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA 18rs21_all.seq
10649 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA coh1_all.seq
10649 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA cjb111_all.seq
10651 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA nem316_all.seq
10650 AAAGTTCTATCATCCTCTTTTGGCGCCTTTTTTCCAGCAGATAATATTAA a909_all.seq

      GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG Majority
               10710     10720     10730     10740     10750
.0694 GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG 2603_all.seq
.0467 GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG 18rs21_all.seq
.0699 GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG coh1_all.seq
.0699 GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG cjb111_all.seq
0701  GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG nem316_all.seq
.0700 GGTAGCTTGGTCTAACAACTCCAGCACTTTATTTACACCACCTATTAATG a909_all.seq

      CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA Majority
               10760     10770     10780     10790     10800
0744  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA 2603_all.seq
0517  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA 18rs21_all.seq
0749  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA coh1_all.seq
0749  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA cjb111_all.seq
0751  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA nem316_all.seq
0750  CAAACTACACCACTCAGATTCAAGCTATTGGGACAACGATTAAGTCACAA a909_all.seq

      ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG Majority
               10810     10820     10830     10840     10850
0794  ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG 2603_all.seq
0567  ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG 18rs21_all.seq
0799  ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG coh1_all.seq
0799  ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG cjb111_all.seq
0801  ATTCCGGAATCGATTTTGACGATTACGGATAAAAAGAGCAGGAAGTTCAG nem316_all.seq
0800  ATTCCGGAATCGATTTTGACGGTTACGGATAAAAAGAGCAGGAAGTTCAG a909_all.seq
```

FIGURE 18 AD

```
      CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA Majority
               10860      10870      10880      10890      10900
10844 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA 2603_ail.seq
10617 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA 18rs21_ail.seq
10849 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA cohl_ail.seq
10849 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA cjb111_ail.seq
10851 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA nem316_ail.seq
10850 CATTAACAAGATTGACGAAGCTAAAGAAGGCTTAGTAGGTGCGACCTTCA a909_ail.seq

      CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT Majority
               10910      10920      10930      10940      10950
10894 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT 2603_ail.seq
10667 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT 18rs21_ail.seq
10899 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT cohl_ail.seq
10899 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT cjb111_ail.seq
10901 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT nem316_ail.seq
10900 CCTTGTCTAAACGCACAACAGTAGCGGCAGATCATCAAGTACAAGGAGAT a909_ail.seq

      TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACCCTTACCTT Majority
               10960      10970      10980      10990      11000
10944 TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACCCTTACCTT 2603_ail.seq
10717 TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACCCTTACCTT 18rs21_ail.seq
10949 TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACCCTTACCTT cohl_ail.seq
10949 TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACTCTTACCTT cjb111_ail.seq
10951 TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACCCTTACCTT nem316_ail.seq
0950  TTCATTCCTGTCAGCAAAGAGACGACAGTCGGTCGGACAACTCTTACCTT a909_ail.seq

      TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA Majority
               11010      11020      11030      11040      11050
0994  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA 2603_ail.seq
0767  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA 18rs21_ail.seq
0999  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA cohl_ail.seq
0999  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA cjb111_ail.seq
1001  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA nem316_ail.seq
1000  TGACAACCTTAAACCTGGATTTTATGACCTTAAAGAAACGAAAGCGCCGA a909_ail.seq

      ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA Majority
               11060      11070      11080      11090      11100
1044  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA 2603_ail.seq
0817  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA 18rs21_ail.seq
1049  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA cohl_ail.seq
1049  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA cjb111_ail.seq
1051  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA nem316_ail.seq
1050  ATGCTTACGTACTTGATCCTAAGACTTATGTTGTGGTCGTTCAAAATTCA a909_ail.seq

      GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC Majority
               11110      11120      11130      11140      11150
1094  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC 2603_ail.seq
0867  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC 18rs21_ail.seq
1099  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC cohl_ail.seq
1099  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC cjb111_ail.seq
1101  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC nem316_ail.seq
1100  GGAAAAACGACAATTGTGGATGAAGCTAACTTCAAAGAGGCTGATTACCC a909_ail.seq

      AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA Majority
               11160      11170      11180      11190      11200
144   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA 2603_ail.seq
917   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA 18rs21_ail.seq
149   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA cohl_ail.seq
149   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA cjb111_ail.seq
151   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA nem316_ail.seq
150   AATGGCTGATAATACCAGCCATGTGGAGTGCCGTAGCGTTGCTACAACGAA a909_ail.seq
```

FIGURE 18 AE

```
      GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT Majority
               11210     11220     11230     11240     11250
11194 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT 2603_ail.seq
10967 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT 18rs21_ail.seq
11199 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT coh1_ail.seq
11199 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT cjb111_ail.seq
11201 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT nem316_ail.seq
11200 GCAAAGGGTAAAAATCCTTTATTTTAAGCACTTTTTCAAGCATTTTGTCT a909_ail.seq

      TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG Majority
               11260     11270     11280     11290     11300
11244 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG 2603_ail.seq
11017 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG 18rs21_ail.seq
11249 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG coh1_ail.seq
11249 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG cjb111_ail.seq
11251 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG nem316_ail.seq
11250 TTATTGAAAAGAGTGATTTTAACATAAAAAAGGTATTAAAAAACATATTG a909_ail.seq

      ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG Majority
               11310     11320     11330     11340     11350
11294 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG 2603_ail.seq
11067 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG 18rs21_ail.seq
11299 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG coh1_ail.seq
11299 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG cjb111_ail.seq
11301 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG nem316_ail.seq
11300 ACGTGACCGTTTGTTTTGAAGTGGCTTGCGTAGACAAAAAAATAGATACG a909_ail.seq

      TCAGATAAATTTCTGGCATTACGAGAACATTTTTAGAGTGTTCTCTTTTT Majority
               11360     11370     11380     11390     11400
11344 TCAGATAAATTTCTGGCATTACGAGAACATTTTTAGAGTGTTCTCTTTTT 2603_ail.seq
11117 TCAGATAAATTTCTGGCATTACGAGAACATTTTTAGAGTGTTCTCTTTTT 18rs21_ail.seq
11349 TCAGATAAATTTCTGGCATTACGAGAACATTTTTAGAGTGTTCTCTTTTT coh1_ail.seq
11349 TCAGATAAATTTCTGGCATTACGAGAACATTTTCTAGAGTGTTCTCTTTTT cjb111_ail.seq
11351 TCAGATAAATTTCTGGCATTACGAGAACATTTTTAGAGTGTTCTCTTTTT nem316_ail.seq
11350 TCAGATAAATTTCTGGCATTACGAGAACATTTTCTAGAGTGTTCTCTTTTT a909_ail.seq

      TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT Majority
               11410     11420     11430     11440     11450
11394 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT 2603_ail.seq
11167 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT 18rs21_ail.seq
11399 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT coh1_ail.seq
11399 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT cjb111_ail.seq
11401 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT nem316_ail.seq
11400 TTAGTTTACGGAGGAAAAATATATATGGAAAAACAGGATTCACGAGTTCT a909_ail.seq

      CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA Majority
               11460     11470     11480     11490     11500
11444 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA 2603_ail.seq
11217 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA 18rs21_ail.seq
11449 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA coh1_ail.seq
11449 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA cjb111_ail.seq
11451 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA nem316_ail.seq
11450 CATCCATTGGGAGGGGAATTCTGGGGACAAGCTCATTGAACACCAAACCA a909_ail.seq

      GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA Majority
               11510     11520     11530     11540     11550
11494 GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA 2603_ail.seq
11267 GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA 18rs21_ail.seq
11499 GCGCAATGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA coh1_ail.seq
11499 GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA cjb111_ail.seq
11501 GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA nem316_ail.seq
11500 GCGCAACGGGGTGGTACTACCAAGTCGATCGTAGCTTTAGTCAACCAAAA a909_ail.seq
```

FIGURE 18 AF

```
      GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT Majority
               11560     11570     11580     11590     11600
11544 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT 2603_ai1.seq
11317 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT 18rs21_ai1.seq
11549 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT coh1_ai1.seq
11549 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT cjb111_ai1.seq
11551 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT nem316_ai1.seq
11550 GGCGAACCGCCCAGAATGATCCAAAGGCACTAGAAAGTGTCCGTAATGAT a909_ai1.seq

      TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC Majority
               11610     11620     11630     11640     11650
11594 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC 2603_ai1.seq
11367 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC 18rs21_ai1.seq
11599 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC coh1_ai1.seq
11599 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC cjb111_ai1.seq
11601 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC nem316_ai1.seq
11600 TCGATTTCGGGCGGTGATGATGTCATGGGTTATGCTTATAGCAAATGTAC a909_ai1.seq

      TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG Majority
               11660     11670     11680     11690     11700
11644 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG 2603_ai1.seq
11417 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG 18rs21_ai1.seq
11649 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG coh1_ai1.seq
11649 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG cjb111_ai1.seq
11651 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG nem316_ai1.seq
11650 TTGGGGAGTTGCGGCACGAATTAATCAGTGGGACTGAAACTCAAAGGTTG a909_ai1.seq

      AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT Majority
               11710     11720     11730     11740     11750
11694 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT 2603_ai1.seq
11467 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT 18rs21_ai1.seq
11699 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT coh1_ai1.seq
11699 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT cjb111_ai1.seq
11701 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT nem316_ai1.seq
11700 AAATGGTGAGAAGATTACCATTACCAGTTCAATGGGAAATGGTCAGGATT a909_ai1.seq

      GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA Majority
               11760     11770     11780     11790     11800
11744 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA 2603_ai1.seq
11517 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA 18rs21_ai1.seq
11749 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA coh1_ai1.seq
11749 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA cjb111_ai1.seq
11751 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA nem316_ai1.seq
11750 GGGTTGGAACAGCCGAAAGACTAGATGGTGAAACTGATACAGTTCCAAAA a909_ai1.seq

      GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG Majority
               11810     11820     11830     11840     11850
11794 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG 2603_ai1.seq
11567 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG 18rs21_ai1.seq
11799 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG coh1_ai1.seq
11799 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG cjb111_ai1.seq
11801 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG nem316_ai1.seq
11800 GAAGGTACTATTCTCTCTTTTTAGGAAAGTAGTTATGGTTCGTATATAGG a909_ai1.seq

      CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA Majority
               11860     11870     11880     11890     11900
11844 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA 2603_ai1.seq
11617 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA 18rs21_ai1.seq
11849 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA coh1_ai1.seq
11849 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA cjb111_ai1.seq
11851 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA nem316_ai1.seq
11850 CTACGGAACTATATCTTTCGTCACATTACATCTACAGATAGTACCATGAA a909_ai1.seq
```

FIGURE 18 AG

Alignment Report of AI-1_alignment, using J. Hein method with Weighted residue weight table.
Thursday, July 29, 2004 5:46 PM


```
      TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT Majority
               11910     11920     11930     11940     11950
11894 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT 2603_all.seq
11667 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT 18rs21_all.seq
11899 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT coh1_all.seq
11899 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT cjb111_all.seq
11901 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT nem316_all.seq
11900 TTTTGCTTATATGACCAAGTAAAGTGAGGATATACTAACAAATGAAATAT a909_all.seq

      TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT Majority
               11960     11970     11980     11990     12000
11944 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT 2603_all.seq
11717 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT 18rs21_all.seq
11949 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT coh1_all.seq
11949 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT cjb111_all.seq
11951 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT nem316_all.seq
11950 TTATTATCGTATTTGTCCATTTTATCGAAAAGTTTGCATATTATCATTAT a909_all.seq

      GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT Majority
               12010     12020     12030     12040     12050
11994 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT 2603_all.seq
11767 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT 18rs21_all.seq
11999 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT coh1_all.seq
11999 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT cjb111_all.seq
12001 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT nem316_all.seq
12000 GTTTGATAAGATGCAAATATAATGATAGTAGGAGCTAAATATGCATATTT a909_all.seq

      AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG Majority
               12060     12070     12080     12090     12100
12044 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG 2603_all.seq
11817 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG 18rs21_all.seq
12049 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG coh1_all.seq
12049 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG cjb111_all.seq
12051 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG nem316_all.seq
12050 AAAAAATCAAGAGTATCCTAAGTGCTTTCCATTTTGAAATTCAAATATAG a909_all.seq

      CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC Majority
               12110     12120     12130     12140     12150
2094  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC 2603_all.seq
1867  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC 18rs21_all.seq
2099  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC coh1_all.seq
2099  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC cjb111_all.seq
2101  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC nem316_all.seq
2100  CTAATAGTTCTAGAACTTCTAATTGTTTTTCGTCGACGATATGAATTTTC a909_all.seq

      AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA Majority
               12160     12170     12180     12190     12200
2144  AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA 2603_all.seq
1917  AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA 18rs21_all.seq
2149  AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA coh1_all.seq
2149  AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA cjb111_all.seq
2151  AATCTTAACTGTTAGAATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA nem316_all.seq
2150  AATCTTAACTGTTAGGATTCCACCTCCCTTTGGTTAAAGAAAAAAGGTCA a909_all.seq

      GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT Majority
               12210     12220     12230     12240     12250
2194  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT 2603_all.seq
1967  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT 18rs21_all.seq
2199  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT coh1_all.seq
2199  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT cjb111_all.seq
2201  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT nem316_all.seq
2200  GGTCGTTTAGATAACTTTGTCAAACAAGCTCAAGCTATCTAAAAATAGTT a909_all.seq
```

FIGURE 18 AH

```
      TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG Majority
             12260     12270     12280     12290     12300
12244 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG 2603_ai1.seq
12017 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG 18rs21_ai1.seq
12249 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG coh1_ai1.seq
12249 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG cjb111_ai1.seq
12251 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG nem316_ai1.seq
12250 TGAAATGGGCATTACTCTAGTTTTTAATAAGCTATCTGATGAGCAGAAGG a909_ai1.seq

      AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT Majority
             12310     12320     12330     12340     12350
12294 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT 2603_ai1.seq
12067 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT 18rs21_ai1.seq
12299 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT coh1_ai1.seq
12299 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT cjb111_ai1.seq
12301 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT nem316_ai1.seq
12300 AGAAGTTAATGCATGTTGGGAAGTCTTATTTTGACTATCAAGAAAATGCT a909_ai1.seq

      CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA Majority
             12360     12370     12380     12390     12400
2344  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA 2603_ai1.seq
2117  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA 18rs21_ai1.seq
2349  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA coh1_ai1.seq
2349  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA cjb111_ai1.seq
2351  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA nem316_ai1.seq
2350  CTTATCCCACAATTAGGTTTTCTATATTCTAAATTAACTAAAAAAATTGA a909_ai1.seq

      ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT Majority
             12410     12420     12430     12440     12450
2394  ACTTGATAATCGGTTGTCTCCGACTGAAAAAAAGTTATTGATTACCTTAT 2603_ai1.seq
2167  ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT 18rs21_ai1.seq
2399  ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT coh1_ai1.seq
2399  ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT cjb111_ai1.seq
2401  ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT nem316_ai1.seq
2400  ACTTGATAATCGGTTGTCTCCGACTGAACAAAAGTTATTGATTACCTTAT a909_ai1.seq

      TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA Majority
             12460     12470     12480     12490     12500
2444  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA 2603_ai1.seq
2217  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA 18rs21_ai1.seq
2449  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA coh1_ai1.seq
2449  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA cjb111_ai1.seq
2451  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA nem316_ai1.seq
2450  TATTACATACTAAAGGTTTAATCATTGATATGTAAGAAGTAAGTCAGCTA a909_ai1.seq

      ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG Majority
             12510     12520     12530     12540     12550
2494  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG 2603_ai1.seq
2267  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG 18rs21_ai1.seq
2499  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG coh1_ai1.seq
2499  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG cjb111_ai1.seq
2501  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG nem316_ai1.seq
2500  ACCGATCTTTCTATTCTAAAACTTATATTGTTGCTTTAGAAATTTTAAAG a909_ai1.seq

      AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA Majority
             12560     12570     12580     12590     12600
2544  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA 2603_ai1.seq
2317  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA 18rs21_ai1.seq
2549  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA coh1_ai1.seq
2549  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA cjb111_ai1.seq
2551  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA nem316_ai1.seq
2550  AGCGTGGATGGCTTCATAATAAACAGAAATCTTACCAATTTGCGAAGCCA a909_ai1.seq
```

FIGURE 18 AI

```
     AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG Majority
             12610     12620     12630     12640     12650
2594 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG 2603_ai1.seq
2367 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG 18rs21_ai1.seq
2599 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG coh1_ai1.seq
2599 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG cjb111_ai1.seq
2601 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG nem316_ai1.seq
2600 AAAAATATGATATTTGAAGAGTCCAAAGATCTAATAGATAGTCCAGTTAG a909_ai1.seq

     AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC Majority
             12660     12670     12680     12690     12700
2644 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC 2603_ai1.seq
2417 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC 18rs21_ai1.seq
2649 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC coh1_ai1.seq
2649 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC cjb111_ai1.seq
2651 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC nem316_ai1.seq
2650 AGAAGCGTTGATTATAAGTGATAAGGATTTTCAAAAATTAAAACAAGAGC a909_ai1.seq

     TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT Majority
             12710     12720     12730     12740     12750
2694 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT 2603_ai1.seq
2467 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT 18rs21_ai1.seq
2699 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT coh1_ai1.seq
2699 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT cjb111_ai1.seq
2701 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT nem316_ai1.seq
2700 TATTATTTTAACCGACTTATTTTAAAGACTTATCATATCTAGGCTTGCTT a909_ai1.seq

     GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA Majority
             12760     12770     12780     12790     12800
2744 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA 2603_ai1.seq
2517 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA 18rs21_ai1.seq
2749 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA coh1_ai1.seq
2749 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA cjb111_ai1.seq
2751 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA nem316_ai1.seq
2750 GATGATTCGGAAAAATACGGAGACTATACTATTTCAAGGAAAAGATACAA a909_ai1.seq

     AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG Majority
             12810     12820     12830     12840     12850
2794 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG 2603_ai1.seq
2567 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG 18rs21_ai1.seq
2799 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG coh1_ai1.seq
2799 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG cjb111_ai1.seq
2801 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG nem316_ai1.seq
2800 AAGTTTCGAATCAAGTCTTCAACTATACATCCTTCAAAGTCATCGGCTAG a909_ai1.seq

     AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC Majority
             12860     12870     12880     12890     12900
2844 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC 2603_ai1.seq
2617 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC 18rs21_ai1.seq
2849 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC coh1_ai1.seq
2849 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC cjb111_ai1.seq
2851 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC nem316_ai1.seq
2850 AGATTTGGAATTATGAACCAATCCCTTTGATTACTAGAAAAATAAATAGC a909_ai1.seq

     TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA Majority
             12910     12920     12930     12940     12950
2894 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA 2603_ai1.seq
2667 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA 18rs21_ai1.seq
2899 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA coh1_ai1.seq
2899 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA cjb111_ai1.seq
2901 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA nem316_ai1.seq
2900 TTGGAGAGGCTAACGTGACACTGGTTGATCCAATCTCGCTTTATTTAACA a909_ai1.seq
```

FIGURE 18 AJ

```
      C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A Majority

                     12960                12970                12980                12990               13000

12944 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A 2603_ai1.seq
12717 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A 18rs21_ai1.seq
12949 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A coh1_ai1.seq
12949 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A cjb111_ai1.seq
12951 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A - nem316_ai1.seq
12950 C T A A G A A T G A T G A A G A C C C T C G T A T T G A A G A A G A A G T T G A G C A G C T A G A A a909_ai1.seq

      G A T A A G A T                                                                              Majority

12994 G A T A A G A T                                                                              2603_ai1.seq
12767 G A T A A G A T                                                                              18rs21_ai1.seq
12999 G A T A A G A T                                                                              coh1_ai1.seq
12999 G A                                                                                          cjb111_ai1.seq
13000 - A                                                                                          nem316_ai1.seq
13000 G A T A A G A T                                                                              a909_ai1.seq
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 18 AK

```
    GGCCTTGTTCCGATGTTGATCCCGATAACTCCTGGCTCATTAATAGCCTG Majority
             10        20        30        40        50
1   GGCCTTGTTCCGATGTTGATTCCGATAACTCCTGGCTCATTAATAGCCTG 2603_ai2.seq
1   GGCCTTGTTCCGATGTTGATTCCGATAACTCCTGGCTCATTAATAGCCTG 18rs21_ai2.seq
1   GGCCTTGTTCCGATGTTGATCCCGATAACTCCTGGCTCATTAATAGCCTG 515_ai2.seq
1   GGCCTTGTTCCGATGTTGATCCCGATAACTCCTGGCTCATTAATAGCCTG cjb111_ai2.seq
1   GGCCTTGTTCCGATGTTGATCCCGATAACTCCTGGCTCATTAATAGCCTG h36b_ai2.seq

    TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT Majority
             60        70        80        90       100
51  TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT 2603_ai2.seq
51  TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT 18rs21_ai2.seq
51  TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT 515_ai2.seq
51  TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT cjb111_ai2.seq
51  TTCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAAT h36b_ai2.seq

    TTTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATG Majority
            110       120       130       140       150
101 TTTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATG 2603_ai2.seq
101 TTTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATG 18rs21_ai2.seq
101 TTTGAAAATAGACTAAGTATTTATTAACTTCAGGCCACTTTCTATGCATG 515_ai2.seq
101 TTTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATG cjb111_ai2.seq
101 TTTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATG h36b_ai2.seq

    AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC Majority
            160       170       180       190       200
151 AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC 2603_ai2.seq
151 AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC 18rs21_ai2.seq
151 AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC 515_ai2.seq
151 AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC cjb111_ai2.seq
151 AAATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAAC h36b_ai2.seq

    TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG Majority
            210       220       230       240       250
201 TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG 2603_ai2.seq
201 TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG 18rs21_ai2.seq
201 TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG 515_ai2.seq
201 TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG cjb111_ai2.seq
201 TATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAG h36b_ai2.seq

    CAACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT Majority
            260       270       280       290       300
251 CAACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT 2603_ai2.seq
251 CAACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT 18rs21_ai2.seq
251 CAACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT 515_ai2.seq
251 CAACTGTTCCATCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT cjb111_ai2.seq
251 CAACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAAT h36b_ai2.seq

    TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG Majority
            310       320       330       340       350
301 TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG 2603_ai2.seq
301 TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG 18rs21_ai2.seq
301 TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG 515_ai2.seq
301 TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG cjb111_ai2.seq
301 TTAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACG h36b_ai2.seq

    ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC Majority
            360       370       380       390       400
351 ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC 2603_ai2.seq
351 ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC 18rs21_ai2.seq
351 ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC 515_ai2.seq
351 ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC cjb111_ai2.seq
351 ATAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACC h36b_ai2.seq

    TCACTACTACAATAAAACTAAAAAAGATTGGAACGTCAGTTAGTCCCAAT Majority
            410       420       430       440       450
401 TCACTACTACAATAAAACTAAAAAAGATTGGAACGTCAGTTAGTCCCAAT 2603_ai2.seq
401 TCACTACTACAATAAAACTAAAAAAGATTGGAACGTCAGTTAGTCCCAAT 18rs21_ai2.seq
401 TCACTACTACAATAAAACTAAAAAAGATTGGAACATCAGTTAGTCCCAAT 515_ai2.seq
401 TCACTACTACAATAAAACTAAAAAAGATTGGAACTAACTGACGTTCCAAT cjb111_ai2.seq
401 TCACTACTACAATAAAACTAAAAAAGATTGGAACGTCAGTTAGTCCCAAT h36b_ai2.seq
```

Figure 19

```
    CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC Majority
             460       470       480       490       500
451 CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC 2603_a12.seq
451 CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC 18rs21_a12.seq
451 CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC 515_a12.seq
451 CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC cjb111_a12.seq
451 CTTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATAC h36b_a12.seq

    GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC Majority
             510       520       530       540       550
501 GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC 2603_a12.seq
501 GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC 18rs21_a12.seq
501 GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC 515_a12.seq
501 GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC cjb111_a12.seq
501 GCAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATC h36b_a12.seq

    TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA Majority
             560       570       580       590       600
551 TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA 2603_a12.seq
551 TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA 18rs21_a12.seq
551 TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA 515_a12.seq
551 TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA cjb111_a12.seq
551 TGTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAA h36b_a12.seq

    TGAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGC Majority
             610       620       630       640       650
601 TGAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGC 2603_a12.seq
601 TGAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGC 18rs21_a12.seq
601 TGAGCAAACCCAAAATACCTTGGCACAACAGTTTCCATATACTCTTAGGC 515_a12.seq
601 TGAGCAAACCCAAAATACCTTGGTACAACAGTTTCCATATACTATTAGGC cjb111_a12.seq
601 TGAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGC h36b_a12.seq

    ATATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCC Majority
             660       670       680       690       700
651 ATATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCC 2603_a12.seq
651 ATATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCC 18rs21_a12.seq
651 ATATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCC 515_a12.seq
651 ATATAGTACTGCAATAAGATAATAATACTCCCAAATATCATAAATGTTCC cjb111_a12.seq
651 ATATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCC h36b_a12.seq

    CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG Majority
             710       720       730       740       750
701 CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG 2603_a12.seq
701 CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG 18rs21_a12.seq
701 CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG 515_a12.seq
701 CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG cjb111_a12.seq
701 CATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGG h36b_a12.seq

    TTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT Majority
             760       770       780       790       800
751 TTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT 2603_a12.seq
751 TTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT 18rs21_a12.seq
751 GTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT 515_a12.seq
751 TTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT cjb111_a12.seq
751 TTAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACAT h36b_a12.seq

    ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA Majority
             810       820       830       840       850
801 ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA 2603_a12.seq
801 ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA 18rs21_a12.seq
801 ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA 515_a12.seq
801 ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA cjb111_a12.seq
801 ATACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCA h36b_a12.seq

    CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC Majority
             860       870       880       890       900
851 CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC 2603_a12.seq
851 CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC 18rs21_a12.seq
851 CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC 515_a12.seq
851 CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC cjb111_a12.seq
851 CTTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGC h36b_a12.seq
```

FIGURE 19A

```
                                      910       920       930       940       950
        TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  Majority
 901    TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  2603_al2.seq
 901    TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  18rs21_al2.seq
 901    TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  515_al2.seq
 901    TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  cjb111_al2.seq
 901    TTACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGA  h36b_al2.seq

                                      960       970       980       990      1000
        TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT  Majority
 951    TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT  2603_al2.seq
 951    TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT  18rs21_al2.seq
 951    TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGCAAT  515_al2.seq
 951    TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT  cjb111_al2.seq
 951    TTTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAAT  h36b_al2.seq

                                     1010      1020      1030      1040      1050
        TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  Majority
1001    TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  2603_al2.seq
1001    TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  18rs21_al2.seq
1001    TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  515_al2.seq
1001    TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  cjb111_al2.seq
1001    TACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGA  h36b_al2.seq

                                     1060      1070      1080      1090      1100
        TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  Majority
1051    TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  2603_al2.seq
1051    TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  18rs21_al2.seq
1051    TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  515_al2.seq
1051    TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  cjb111_al2.seq
1051    TAATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATT  h36b_al2.seq

                                     1110      1120      1130      1140      1150
        AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  Majority
1101    AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  2603_al2.seq
1101    AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  18rs21_al2.seq
1101    AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  515_al2.seq
1101    AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  cjb111_al2.seq
1101    AGTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGAT  h36b_al2.seq

                                     1160      1170      1180      1190      1200
        AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  Majority
1151    AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  2603_al2.seq
1151    AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  18rs21_al2.seq
1151    AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  515_al2.seq
1151    AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  cjb111_al2.seq
1151    AGAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACC  h36b_al2.seq

                                     1210      1220      1230      1240      1250
        ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  Majority
1201    ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  2603_al2.seq
1201    ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  18rs21_al2.seq
1201    ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  515_al2.seq
1201    ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  cjb111_al2.seq
1201    ATTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTT  h36b_al2.seq

                                     1260      1270      1280      1290      1300
        AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  Majority
1251    AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  2603_al2.seq
1251    AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  18rs21_al2.seq
1251    AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  515_al2.seq
1251    AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  cjb111_al2.seq
1251    AGCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCA  h36b_al2.seq

                                     1310      1320      1330      1340      1350
        TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  Majority
1301    TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  2603_al2.seq
1301    TAGATACAGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  18rs21_al2.seq
1301    TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  515_al2.seq
1301    TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  cjb111_al2.seq
1301    TAGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCT  h36b_al2.seq
```

FIGURE 19B

```
     TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA Majority
              1360      1370      1380      1390      1400
1351 TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA 2603_a12.seq
1351 TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA 18rs21_a12.seq
1351 TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA 515_a12.seq
1351 TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA cjb111_a12.seq
1351 TTTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAA h36b_a12.seq

     AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC Majority
              1410      1420      1430      1440      1450
1401 AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC 2603_a12.seq
1401 AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC 18rs21_a12.seq
1401 AGCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCAC 515_a12.seq
1401 AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC cjb111_a12.seq
1401 AGCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCAC h36b_a12.seq

     GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA Majority
              1460      1470      1480      1490      1500
1451 GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA 2603_a12.seq
1451 GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA 18rs21_a12.seq
1451 GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA 515_a12.seq
1451 GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA cjb111_a12.seq
1451 GATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAA h36b_a12.seq

     CTTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA Majority
              1510      1520      1530      1540      1550
1501 CTTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA 2603_a12.seq
1501 CTTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA 18rs21_a12.seq
1501 CTTGATTTCAAATCAAATAAAATAAAAGTAACTAACATCGGAAGGATTGA 515_a12.seq
1501 CTTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA cjb111_a12.seq
1501 CTTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGA h36b_a12.seq

     AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA Majority
              1560      1570      1580      1590      1600
1551 AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA 2603_a12.seq
1551 AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA 18rs21_a12.seq
1551 AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA 515_a12.seq
1551 AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCA cjb111_a12.seq
1551 AAAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAAGCA h36b_a12.seq

     TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA Majority
              1610      1620      1630      1640      1650
1601 TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA 2603_a12.seq
1601 TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA 18rs21_a12.seq
1601 TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA 515_a12.seq
1601 TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA cjb111_a12.seq
1601 TCATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTA h36b_a12.seq

     CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG Majority
              1660      1670      1680      1690      1700
1651 CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG 2603_a12.seq
1651 CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG 18rs21_a12.seq
1651 CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG 515_a12.seq
1651 CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG cjb111_a12.seq
1651 CGTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAG h36b_a12.seq

     CTCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATC Majority
              1710      1720      1730      1740      1750
1701 CTCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATC 2603_a12.seq
1701 CTCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATC 18rs21_a12.seq
1701 CTCCAAAGGTAAGCGTAGGTACGCGAAAAAAACCTTTGTCTTCTCCCATC 515_a12.seq
1701 CTCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATC cjb111_a12.seq
1701 CTCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATC h36b_a12.seq

     CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG Majority
              1760      1770      1780      1790      1800
1750 CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG 2603_a12.seq
1750 CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG 18rs21_a12.seq
1751 CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG 515_a12.seq
1750 CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG cjb111_a12.seq
1750 CAGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCG h36b_a12.seq
```

FIGURE 19C

```
     GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC Majority
              1810      1820      1830      1840      1850
1800 GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC 2603_a12.seq
1800 GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC 18rs21_a12.seq
1801 GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC 515_a12.seq
1800 GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC cjb111_a12.seq
1800 GACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTC h36b_a12.seq

     GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT Majority
              1860      1870      1880      1890      1900
1850 GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT 2603_a12.seq
1850 GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT 18rs21_a12.seq
1851 GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT 515_a12.seq
1850 GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT cjb111_a12.seq
1850 GGGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTT h36b_a12.seq

     GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT Majority
              1910      1920      1930      1940      1950
1900 GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT 2603_a12.seq
1900 GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT 18rs21_a12.seq
1901 GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT 515_a12.seq
1900 GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT cjb111_a12.seq
1900 GCAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCAT h36b_a12.seq

     ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT Majority
              1960      1970      1980      1990      2000
1950 ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT 2603_a12.seq
1950 ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT 18rs21_a12.seq
1951 ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT 515_a12.seq
1950 ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT cjb111_a12.seq
1950 ATTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTT h36b_a12.seq

     TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA Majority
              2010      2020      2030      2040      2050
2000 TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA 2603_a12.seq
2000 TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA 18rs21_a12.seq
2001 TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA 515_a12.seq
2000 TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA cjb111_a12.seq
2000 TAGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGA h36b_a12.seq

     ATTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTG Majority
              2060      2070      2080      2090      2100
2050 ATTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTG 2603_a12.seq
2050 ATTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTG 18rs21_a12.seq
2051 ATTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTG 515_a12.seq
2050 ATTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTG cjb111_a12.seq
2050 ATTGTCAGTCCTCCGCCAATAATCATTCCTGTTTTAGGAAGAAATGATTG h36b_a12.seq

     TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT Majority
              2110      2120      2130      2140      2150
2100 TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT 2603_a12.seq
2100 TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT 18rs21_a12.seq
2101 TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT 515_a12.seq
2100 TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT cjb111_a12.seq
2100 TGGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTT h36b_a12.seq

     TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT Majority
              2160      2170      2180      2190      2200
2150 TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT 2603_a12.seq
2150 TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT 18rs21_a12.seq
2151 TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT 515_a12.seq
2150 TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT cjb111_a12.seq
2150 TTTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACT h36b_a12.seq

     ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCG Majority
              2210      2220      2230      2240      2250
2200 ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCG 2603_a12.seq
2200 ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCG 18rs21_a12.seq
2201 ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTACTTCTCG 515_a12.seq
2200 ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCG cjb111_a12.seq
2200 ACAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCG h36b_a12.seq
```

FIGURE 19D

```
      A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  Majority
                2260                2270                2280                2290                2300
2250  A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  2603_al2.seq
2250  A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  18rs21_al2.seq
2251  A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  515_al2.seq
2250  A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  cjb111_al2.seq
2250  A A A A A T A T A C T T A C C A G G T A A T A A A C C T T C A A C C T C A A T T T C T C C C T T A T  h36b_al2.seq

      C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  Majority
                2310                2320                2330                2340                2350
2300  C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  2603_al2.seq
2300  C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  18rs21_al2.seq
2301  C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  515_al2.seq
2300  C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  cjb111_al2.seq
2300  C A T C A G T T A C T A A T G A A G T A A T C C C A T C T T G A T C G G T C G T A A A T C G T C C A  h36b_al2.seq

      T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  Majority
                2360                2370                2380                2390                2400
2350  T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  2603_al2.seq
2350  T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  18rs21_al2.seq
2351  T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  515_al2.seq
2350  T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  cjb111_al2.seq
2350  T T T T T A A A G C G A A C T G G C T G A T T C T G G T T A T C G T A T A A T A C A A A T A T T A C  h36b_al2.seq

      T C C G G A T A G C C T T T T C T T T A T C T T T C C T T C T T T T G T A T A T T T A A T A A G T T  Majority
                2410                2420                2430                2440                2450
2400  T C C G G A T A G C C T T T T C T T T A T C T T T C C T T C T T T T G T A T A T T T A A T A A G T T  2603_al2.seq
2400  T C C G G A T A G C C T T T T C T T T A T C T T T C C T T C T T T T G T A T A T T T A A T A A G T T  18rs21_al2.seq
2401  T C C T G A T A G C C T T T T C T T T A T C T T T C C T T C T T T T G T A T A T T T A A T A A G T T  515_al2.seq
2400  T C C G G A T A G C C T T T T C T T T A T C T T T C C T T C T T T T G T A T A T T T A A T A A G T T  cjb111_al2.seq
2400  T C C G G A T A G C C Y T T T C T T T A T C T T T C C T T C T T T T G K A T A T T T A A T A A G T T  h36b_al2.seq

      T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  Majority
                2460                2470                2480                2490                2500
2450  T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  2603_al2.seq
2450  T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  18rs21_al2.seq
2451  T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  515_al2.seq
2450  T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  cjb111_al2.seq
2450  T T A A T C G G C C T G T T T C A A C T T T T C G C T T A G G A T T T A T C T G T A A T T G A T T T  h36b_al2.seq

      G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  Majority
                2510                2520                2530                2540                2550
2500  G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  2603_al2.seq
2500  G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  18rs21_al2.seq
2501  G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  515_al2.seq
2500  G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  cjb111_al2.seq
2500  G A T A A C T T A T C A T C T G G T A A T T C A A T A T A A A A A G G T A C T A T T G T T G A A A C  h36b_al2.seq

      G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  Majority
                2560                2570                2580                2590                2600
2550  G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  2603_al2.seq
2550  G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  18rs21_al2.seq
2551  G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  515_al2.seq
2550  G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  cjb111_al2.seq
2550  G C T T T G A T C A G C T T T A T A A G C A C G A C C A A A G T A C G A A C C A T T T G G G A G T G  h36b_al2.seq

      C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  Majority
                2610                2620                2630                2640                2650
2600  C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  2603_al2.seq
2600  C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  18rs21_al2.seq
2601  C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  515_al2.seq
2600  C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  cjb111_al2.seq
2600  C T A T C T T T G T C T G A C C A T T A G T A T C A G T A G G A G A A G T C A A G A T A C T C T T A  h36b_al2.seq

      T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  Majority
                2660                2670                2680                2690                2700
2650  T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  2603_al2.seq
2650  T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  18rs21_al2.seq
2651  T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  515_al2.seq
2650  T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  cjb111_al2.seq
2650  T A C T T C T G G T T C A A T T C G C T A T C T G T C A T T T G G C T C A A T A A A T C A A C T T T  h36b_al2.seq
```

FIGURE 19E

```
     TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG Majority
              2710      2720      2730      2740      2750
2700 TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG 2603_al2.seq
2700 TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG 18rs21_al2.seq
2701 TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG 515_al2.seq
2700 TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG cjb111_al2.seq
2700 TAAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAG h36b_al2.seq

     TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA Majority
              2760      2770      2780      2790      2800
2750 TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA 2603_al2.seq
2750 TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA 18rs21_al2.seq
2751 TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA 515_al2.seq
2750 TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA cjb111_al2.seq
2750 TATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCA h36b_al2.seq

     GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA Majority
              2810      2820      2830      2840      2850
2800 GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA 2603_al2.seq
2800 GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA 18rs21_al2.seq
2801 GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA 515_al2.seq
2800 GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA cjb111_al2.seq
2800 GCATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAA h36b_al2.seq

     GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT Majority
              2860      2870      2880      2890      2900
2850 GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT 2603_al2.seq
2850 GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT 18rs21_al2.seq
2851 GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT 515_al2.seq
2850 GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT cjb111_al2.seq
2850 GCAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATT h36b_al2.seq

     TCCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG Majority
              2910      2920      2930      2940      2950
2900 TCCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG 2603_al2.seq
2900 TCCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG 18rs21_al2.seq
2901 TCCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG 515_al2.seq
2900 TCCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG cjb111_al2.seq
2900 TCCCTTTCTTTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAG h36b_al2.seq

     CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG Majority
              2960      2970      2980      2990      3000
2950 CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG 2603_al2.seq
2950 CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG 18rs21_al2.seq
2951 CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG 515_al2.seq
2950 CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG cjb111_al2.seq
2950 CGACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAG h36b_al2.seq

     TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC Majority
              3010      3020      3030      3040      3050
3000 TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC 2603_al2.seq
3000 TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC 18rs21_al2.seq
3001 TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC 515_al2.seq
3000 TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC cjb111_al2.seq
3000 TGTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTAC h36b_al2.seq

     ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT Majority
              3060      3070      3080      3090      3100
3050 ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT 2603_al2.seq
3050 ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT 18rs21_al2.seq
3051 ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT 515_al2.seq
3050 ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT cjb111_al2.seq
3050 ATACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAAT h36b_al2.seq

     AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT Majority
              3110      3120      3130      3140      3150
3100 AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT 2603_al2.seq
3100 AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT 18rs21_al2.seq
3101 AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT 515_al2.seq
3100 AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT cjb111_al2.seq
3100 AAGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAAT h36b_al2.seq
```

FIGURE 19F

```
     AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA Majority
              3160      3170      3180      3190      3200
3150 AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA 2603_a12.seq
3150 AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA 18rs21_a12.seq
3151 AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA 515_a12.seq
3150 AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA cjb111_a12.seq
3150 AATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATA h36b_a12.seq

     GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA Majority
              3210      3220      3230      3240      3250
3200 GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA 2603_a12.seq
3200 GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA 18rs21_a12.seq
3201 GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA 515_a12.seq
3200 GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA cjb111_a12.seq
3200 GGCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAA h36b_a12.seq

     GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT Majority
              3260      3270      3280      3290      3300
3250 GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT 2603_a12.seq
3250 GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT 18rs21_a12.seq
3251 GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT 515_a12.seq
3250 GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT cjb111_a12.seq
3250 GTTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCT h36b_a12.seq

     GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTC Majority
              3310      3320      3330      3340      3350
3300 GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTC 2603_a12.seq
3300 GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTC 18rs21_a12.seq
3301 GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAAGAAGTTCCTTC 515_a12.seq
3300 GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTC cjb111_a12.seq
3300 GTGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAAGAAGTTCCTTC h36b_a12.seq

     AAGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA Majority
              3360      3370      3380      3390      3400
3350 AAGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA 2603_a12.seq
3350 AAGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA 18rs21_a12.seq
3351 AAGGTGTCCTGATCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA 515_a12.seq
3350 AAGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA cjb111_a12.seq
3350 AAGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGA h36b_a12.seq

     TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA Majority
              3410      3420      3430      3440      3450
3400 TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA 2603_a12.seq
3400 TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA 18rs21_a12.seq
3401 TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA 515_a12.seq
3400 TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA cjb111_a12.seq
3400 TAGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCA h36b_a12.seq

     GCAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC Majority
              3460      3470      3480      3490      3500
3450 GCAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC 2603_a12.seq
3450 GCAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC 18rs21_a12.seq
3451 GTAATCTCAAGCATATGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC 18rs21_a12.seq
3450 GCAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC 515_a12.seq
3450 GCAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTC cjb111_a12.seq
     h36b_a12.seq

     AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT Majority
              3510      3520      3530      3540      3550
3500 AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT 2603_a12.seq
3500 AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT 18rs21_a12.seq
3501 AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT 515_a12.seq
3500 AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT cjb111_a12.seq
3500 AGTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTT h36b_a12.seq

     TTGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTT Majority
              3560      3570      3580      3590      3600
3550 TTGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTT 2603_a12.seq
3550 TTGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTT 18rs21_a12.seq
3551 TTGCTAACTCAAATCGTCTATTAATCTCTTTAGCATTTAATTTTTGGGTT 515_a12.seq
3550 TTGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTT cjb111_a12.seq
3550 TTGCTAACTCAAATCGTCTATTAATCTCTTTAGCATTTAATTTTTGGGTT h36b_a12.seq
```

FIGURE 19G

```
     TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA Majority
              3610      3620      3630      3640      3650
3600 TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA 2603_a12.seq
3600 TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA 18rs21_a12.seq
3601 TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA 515_a12.seq
3600 TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA cjb111_a12.seq
3600 TGATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCA h36b_a12.seq

     ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATA Majority
              3660      3670      3680      3690      3700
3650 ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATA 2603_a12.seq
3650 ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATA 18rs21_a12.seq
3651 ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCTACTAAGAATAATA 515_a12.seq
3650 ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATA cjb111_a12.seq
3650 ATTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATA h36b_a12.seq

     GTATCAAGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG Majority
              3710      3720      3730      3740      3750
3700 GTATCAAGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG 2603_a12.seq
3700 GTATCAAGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG 18rs21_a12.seq
3701 ATATCAAACCTATATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG 515_a12.seq
3700 GTATCAGGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG cjb111_a12.seq
3700 GTATCAAGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGG h36b_a12.seq

     TTTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTT Majority
              3760      3770      3780      3790      3800
3750 TTTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTT 2603_a12.seq
3750 TTTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTT 18rs21_a12.seq
3751 TTTTTCTGAAATTTTCCTCGCATTATGATTCAATTCCTTTTCTAACACTT 515_a12.seq
3750 TTTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTT cjb111_a12.seq
3750 TTTTTCTGAAATCTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTT h36b_a12.seq

     GCTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC Majority
              3810      3820      3830      3840      3850
3800 GCTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC 2603_a12.seq
3800 GCTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC 18rs21_a12.seq
3801 GATAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC 515_a12.seq
3800 GCTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC cjb111_a12.seq
3800 GCTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGC h36b_a12.seq

     AATAATAACTAAAGATATATAGAAAAGATATCTATAAATCGTGTTTAAAT Majority
              3860      3870      3880      3890      3900
3850 AATAATAACTAAAGATATATAGAATAGATATCTATAAATCGTGTTTAAAT 2603_a12.seq
3850 AATAATAACTAAAGATATATAGAATAGATATCTATAAATCGTGTTTAAAT 18rs21_a12.seq
3851 AATAATAACTAAAGATATATAGAAAAGATATCTATAAATCGTGTTTAAAT 515_a12.seq
3850 AATAATAACTAAAGATATATAGAAAAGATATCTATAAATCGTGTTTAAAT cjb111_a12.seq
3850 AATAATAACTAAAGATATATAGAAAAGATATCTATAAATCGTGTTTAAAT h36b_a12.seq

     GACCGTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGA Majority
              3910      3920      3930      3940      3950
3900 GACCGTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGA 2603_a12.seq
3900 GACCGTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGA 18rs21_a12.seq
3901 GACCGTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGA 515_a12.seq
3900 GACCTTCTTTCATTAATTTTTCATCAATAAGACCTTTATAAGGAATACGA cjb111_a12.seq
3900 GACCTTCTTTCATTAATTTTTCATCAATAGGACCTTTATAAGGAATACGA h36b_a12.seq

     TGTCCCCTTACTAAAAGTCTGTGTGTATTGACCATAATCGGGGTGCAAGT Majority
              3960      3970      3980      3990      4000
3950 TGTCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGT 2603_a12.seq
3950 TGTCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGT 18rs21_a12.seq
3951 TGACCCCTTACTAAAAGTCTGTGTGTATTGACCATAATCGGGGTGCAAGT 515_a12.seq
3950 TGTCCCCTTACTAAAAGTCTGTGTGTATTGACCATAATCGGGGTGCAAGT cjb111_a12.seq
3950 TGTCCCCTTACTAAAAGTCTGTGTGTATTGACCATAATCGGGGTGCAAGT h36b_a12.seq

     CAATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGT Majority
              4010      4020      4030      4040      4050
4000 CAATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGT 2603_a12.seq
4000 CAATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGT 18rs21_a12.seq
4001 CAATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGT 515_a12.seq
4000 TAATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGT cjb111_a12.seq
4000 CAATAAGGTTGCATAATCATGTCCAGGAACAACCAATAAATCTGAAAAGT h36b_a12.seq
```

FIGURE 19H

```
     TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT Majority
          4060      4070      4080      4090      4100
4050 TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT 2603_a12.seq
4050 TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT 18rs21_a12.seq
4051 TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT 515_a12.seq
4050 TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT cjb111_a12.seq
4050 TATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCT h36b_a12.seq

     TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT Majority
          4110      4120      4130      4140      4150
4100 TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT 2603_a12.seq
4100 TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT 18rs21_a12.seq
4101 TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT 515_a12.seq
4100 TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT cjb111_a12.seq
4100 TTTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTT h36b_a12.seq

     AGAAAAGAGTTCTTTATCTGGAATTCCTGAATGCGCTGTTATAACAGTAT Majority
          4160      4170      4180      4190      4200
4150 AGAAAAGAGTTCTTTATCTGGAATTCCTGAATGCGCTGTTATAACAGTAT 2603_a12.seq
4150 AGAAAAGAGTTCTTTATCTGGAATTCCTGAATGCGCTGTTATAACAGTAT 18rs21_a12.seq
4151 AGAAAAGAGCTCTTTATCTGGAATTCCTGAGTGCGCTGTTATAACGGTAT 515_a12.seq
4150 AGAAAAGAGTTCTTTATCTGGAATTCCTGAGTGCGCTGTTATAACGGTAT cjb111_a12.seq
4150 AGAAAAGAGTTCTTTATCTGGAATTCCTGAATGTGCTGTTATAACAGTAT h36b_a12.seq

     GTGTACTATTGCCTCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCT Majority
          4210      4220      4230      4240      4250
4200 GTGTACTATTGCCCCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCT 2603_a12.seq
4200 GTGTACTATTGCCCCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCT 18rs21_a12.seq
4201 GGGTGCTATTGCCTCCAATTGGAAGAGAGGTACCTTCTAAATGACCTGCT 515_a12.seq
4200 GTGTGCTATTTCCTCCAATTGGAAGAGAGAGTACCTTCTAAATGCCCTGCT cjb111_a12.seq
4200 GTGTACTATTGCCTCCAATTGGAAGAGAGGTGCCTTCTAAATGCCCTGCT h36b_a12.seq

     CCTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTG Majority
          4260      4270      4280      4290      4300
4250 CCTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTG 2603_a12.seq
4250 CCTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTG 18rs21_a12.seq
4251 CCTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTG 515_a12.seq
4250 CCTTTAGATAGAACTTTTTGACTTGAACCTGCAAATATAGGGAGTTTTTG cjb111_a12.seq
4250 CCTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTG h36b_a12.seq

     ACCTATCTTAGGAACTGAAATTGTGCCGATTTTTTCACTTACCTCTAACA Majority
          4310      4320      4330      4340      4350
4300 ACCTATCTTAGGAACTGAAATTGTTCCGATTTTTTCACTTACCTCTAACA 2603_a12.seq
4300 ACCTATCTTAGGAACTGAAATTGTTCCGATTTTTTCACTTACCTCTAACA 18rs21_a12.seq
4301 ACCTATCTTAGGAACTGAAATTATGCCGATTTTTTCACTTACCTCTAACA 515_a12.seq
4300 ACCTATCTTAGGAACTGAAATTGTGCCGATTTTTTCACTTACCTCTAACA cjb111_a12.seq
4300 ACCTATCTTAGGAACTGAAATTGTGCCGATTTTTTCACTTACCTCTAACA h36b_a12.seq

     TACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGA Majority
          4360      4370      4380      4390      4400
4350 TACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGA 2603_a12.seq
4350 TACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGA 18rs21_a12.seq
4351 TACGGGCGTACTCTGCTATCCCCTTTTGAATTCGTTTTTTCTCATAAGGA 515_a12.seq
4350 TACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTCTTCTCATAAGGA cjb111_a12.seq
4350 TACGGGCGTACTCTGCTACCCCTTTTTGAATTCGTTTTTTCTCATAAGGA h36b_a12.seq

     TCTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGC Majority
          4410      4420      4430      4440      4450
4400 TCTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGC 2603_a12.seq
4400 TCTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGC 18rs21_a12.seq
4401 TCTTCAAGATGGACATTATTTAAAGAATCATTATATGCTTGTGCTAGAGC 515_a12.seq
4400 TCTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGT cjb111_a12.seq
4400 TCTTCAAAATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGC h36b_a12.seq

     CATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA Majority
          4460      4470      4480      4490      4500
4450 CATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA 2603_a12.seq
4450 CATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA 18rs21_a12.seq
4451 CATACGCCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA 515_a12.seq
4450 CATACGACGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA cjb111_a12.seq
4450 CATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAA h36b_a12.seq
```

FIGURE 19I

```
     AATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACC Majority
              4510      4520      4530      4540      4550
4500 AATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACC 2603_ai2.seq
4500 AATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACC 18rs21_ai2.seq
4501 AATCTTGTGTTTGGTTATTAGATTCTATCGTATAGTAAAATCGTGATACC 515_ai2.seq
4500 AATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACC cjb111_ai2.seq
4500 AATCCTGTGTTTGATTATTGGATTCTATCGTATAGTAAAAACGTGATACC h36b_ai2.seq

     ACTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGG Majority
              4560      4570      4580      4590      4600
4550 ACTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGG 2603_ai2.seq
4550 ACTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGG 18rs21_ai2.seq
4551 ACTGGATACAATAAGATAGATAGACCTATTAGAAAAAGAATGATAAAAGG 515_ai2.seq
4550 ACTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGG cjb111_ai2.seq
4550 ACTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGG h36b_ai2.seq

     AAGATTTGACTTCTTCTTTTT---TGTTTTTTTGATGATTTTTTTAGTCT Majority
              4610      4620      4630      4640      4650
4600 AAGATTTGACTTCTTCTTTTTTTTTGTTTTTTTGATGATTTTTTTAGTCT 2603_ai2.seq
4600 AAGATTTGACTTCTTCTTTTTTTTTGTTTTTTTGATGATTTTTTTAGTCT 18rs21_ai2.seq
4601 CAGATTTGATTTCTTTTTTTT---TGTTTTTTTGATAATTTTTTTAGTCT 515_ai2.seq
4600 AAGATTTGACTTCTTCTTTTT---AGTTTTTTTGTTGATATTTTTTAGTCT cjb111_ai2.seq
4600 AAGATTTGACTTCTTCTTTTT---AGTTTTTTTGATGATATTTTTTAGTCT h36b_ai2.seq

     TCACGTCATCTCCTAAATAATGGCTCTTGCTTATGATCTAAGAGTACTTC Majority
              4660      4670      4680      4690      4700
4650 TCACGTCATCTCCTAGATAATGGCTCTTGCTTATGATCTAAGAGTACTTC 2603_ai2.seq
4650 TCACGTCATCTCCTAGATAATGGCTCTTGCTTATGATCTAAGAGTACTTC 18rs21_ai2.seq
4648 TCACGTCATCTCCTAAATAATGGCTCTTGCTTATGATCTAAGAGTACTTC 515_ai2.seq
4647 TCACGTCATCTCCTAAATAATGGCTCTTGCTTATGATCTAAGAGTACTTC cjb111_ai2.seq
4647 TCACGTCATCTCCTAAATAATGGCTCTTGCTTATGATCTAAGAGTACTTC h36b_ai2.seq

     TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACAT Majority
              4710      4720      4730      4740      4750
4700 TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGCGCTTATACAT 2603_ai2.seq
4700 TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGCGCTTATACAT 18rs21_ai2.seq
4698 TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTACACAT 515_ai2.seq
4697 TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACAT cjb111_ai2.seq
4697 TACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACAT h36b_ai2.seq

     CATCAAAGACTAGCCTTAAGCTTCCTCTGATTGACGTCTTTTCATGATAA Majority
              4760      4770      4780      4790      4800
4750 CATCAAAGACTAGCCTTAAGCTTCCTTTGATTGGCGTCTTTTCATGATAA 2603_ai2.seq
4750 CATCAAAGACTAGCCTTAAGCTTCCTTTGATTGGCGTCTTTTCATGATAA 18rs21_ai2.seq
4748 CATCAAAAGCTAATCTTAAACTTCCTCTGATTGACGTCTTTTCATGATAA 515_ai2.seq
4747 CATCAAAGACTAGCCTTAAGCTTCCTCTGATTGACGTTTTTTCATGATAA cjb111_ai2.seq
4747 CATCAAAGACTAGCCTTAAGCTTCCTCTGATTGACGTCTTTTCATGATAA h36b_ai2.seq

     CTACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTA Majority
              4810      4820      4830      4840      4850
4800 CTACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTA 2603_ai2.seq
4800 CTACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTA 18rs21_ai2.seq
4798 CTACCGCTCCAAGCATAATACTTAATCCAATAATTGTGAAAAAAATTGTA 515_ai2.seq
4797 CTACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTA cjb111_ai2.seq
4797 TTACCGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTA h36b_ai2.seq

     CCAATACCACCTGTTTGTGGGATTGTTACCTTTTTATTTTCTACTCGTTG Majority
              4860      4870      4880      4890      4900
4850 CCAATACCACCTGTTTGTGGGATTGTTACCTTTTTATTTTCTACACGTGT 2603_ai2.seq
4850 CCAATACCACCTGTTTGTGGGATTGTTACCTTTTTATTTTCTACACGTGT 18rs21_ai2.seq
4848 CCAATACCACCTGTTTGTGGAATAGTAACCTTCTTATTGATAACTTGTTG 515_ai2.seq
4847 CCAATACCACCTGTTTGTGGGATGGTTACTTTTTTGTTTTGAACTTGTTG cjb111_ai2.seq
4847 CCAATACCACCTGTTTGTGGAATAGTCACTTTTTTGTTTTCTATGCGTTG h36b_ai2.seq

     CGCATCTTTTTTTTTGCTGCTAGCAGCGTAGTCAATGTTACCTG---AAC Majority
              4910      4920      4930      4940      4950
4900 CGCATCTTTTTTGGTTGCTGTTAGCAACGTAGTCAATGTTACC------AC 2603_ai2.seq
4900 CGCATCTTTTTTGGTTGCTGTTAGCAACGTAGTCAATGTTACC------AC 18rs21_ai2.seq
4898 TGCATCTTTAGTTTTAGAACCTTGGGTATAGTCAATATCTTGAGCAGAAC 515_ai2.seq
4897 GGCATCTTTTTTTACAGAGCCTTTATCATATGCGATGTCAGTTGTAGCTC cjb111_ai2.seq
4897 AGCGTCTTTTTTTATTGCTGCTAGCAGTGTAATCAATGTTACCTG---AGT h36b_ai2.seq
```

FIGURE 19J

```
     CTTTTCTGTATGACCTTTGATTAACTTCAAACTTAACATCACCTGCCAAT Majority
              4960      4970      4980      4990      5000
4944 CTGTTATGTATGACCCTTGATTAACTACAAACTTAATATTACCTGCCAAC 2603_a12.seq
4944 CTGTTATGTATGACCCTTGATTAACTACAAACTTAATATTACCTGCCAAC 18rs21_a12.seq
4948 CTTTACTGTATGAAGTAGCATTAACATTAAACGAAACATCACCTGATAAT 515_a12.seq
4947 CTTTGCTATATGAGTGGCAGTTACTTCAAAGTTTACATCACCTGACAAT cjb111_a12.seq
4944 CGTCTGCCTTTGAGCTGTTTCCAACTTTGAAATCAACATCACCTGCTAAT h36b_a12.seq

     TTAGCATAACCTGCTGGAGCTTGTGTTTCTTCAAGGTTGTAAGTACCTTC Majority
              5010      5020      5030      5040      5050
4994 TTAGCAAATCCTGCTGGAGCAAGTGTTTCTTCAAGGTTGTAAGTACCGTC 2603_a12.seq
4994 TTAGCAAATCCTGCTGGAGCAAGTGTTTCTTCAAGGTTGTAAGTACCGTC 18rs21_a12.seq
4998 TTAGCATAACCAGTTGGTGCTTGTGTTTCTTCAATGAGTATTGACCTTC 515_a12.seq
4997 GTCGCATAACCTGCTGGTGCTTGAGTTTCTTCCAAGCCATAAGTGCCTTT cjb111_a12.seq
4994 TTAGCGTAACCAGCTGGAGCTTGTGTTTCTTCAAGTTTATAATCACCAGT h36b_a12.seq

     TGCAAGACCTGTAACTTCAAATTGACCTTGATCGTTTGAAGTCAATTTAA Majority
              5060      5070      5080      5090      5100
5044 TGCAAGACCTGTAACTTCAAATTGACCTTGATCGTTTGAAGTGTAGGTAA 2603_a12.seq
5044 TGCAAGACCTGTAACTTCAAATTGACCTTGATCGTTTGAAGTGTAGGTAA 18rs21_a12.seq
5048 AGTCAAGCCAGTAATTTCAAATTGACCTTTATCGTTAGAAATCAATTTAA 515_a12.seq
5047 ATCAAACCAGTAATTTCAAATTGACCACCGGCGTTAGAGATCAATTTAA cjb111_a12.seq
5044 TGCAAGACCGCTAACTTCAAATTGACCGTTCAGCGTTTGAAGTCAGTTTAA h36b_a12.seq

     CAACTTTAGCAGC-TTTTTATCTGTTACCCACTCATAAGCTGTACGAGCT Majority
              5110      5120      5130      5140      5150
5094 TGGCTCTAGC-----TTATCTGTTATCCACTCATAAGCTGTACGAGCC 2603_a12.seq
5094 TGGCTCTAGC-----TTATCTGTTATCCACTCATAAGCTGTACGAGCC 18rs21_a12.seq
5098 CAACATTCTTAGCAATTTTTATCTCAAACCCATTCGTAGGAGTAGTTAGCT 515_a12.seq
5097 CAACATTATCAGCATTTTTATCTGCAACCCATTCATAGGAGTAGTTAGCT cjb111_a12.seq
5094 CAACGTTAGCATCTTGTTTATTAGTTACCCACTCGTAAGCTGTACGAGCT h36b_a12.seq

     TCAATGAAGGCTGCATTGTAAGCTTCTTGTTTAGTTTTGATAGTTGCTTT Majority
              5160      5170      5180      5190      5200
5138 TCAATGAAGGCTGCATCGTAATCTGCTTGTTTAGTTTTGATAAGTTCTTT 2603_a12.seq
5138 TCAATGAAGGCTGCATCGTAATCTGCTTGTTTAGTTTTGATAAGTTCTTT 18rs21_a12.seq
5148 TTAACAAAGGCATCATTGTAAGCTTTTTGTTTTTCACTAAGGGTAGCCAA 515_a12.seq
5147 TTAACAAAAGCGTCATTGTAAGCTTTTTGTTTTTGATCAAGAGTAGCCAA cjb111_a12.seq
5144 ACGATGAAGGCTGCATTGTAAGCGTCTTGAGCAGTTTTGATTTCTGCTTT h36b_a12.seq

     TGCTGTTTTACCTTTTTGACCTTTTTGTTCTTCTGCAGACAACTTGTTAT Majority
              5210      5220      5230      5240      5250
5188 TGCAGTAATTCCTTTTTCACCTTTTTGGTCTGGTGCAGACAACTTGTTAT 2603_a12.seq
5188 TGCAGTAATTCCTTTTTCACCTTTTTGGTCTGGTGCAGACAACTTGTTAT 18rs21_a12.seq
5198 TGCTGATTTACCATCTTGACCTTCTTGTTTTTCTTTAGTCAAATCATTGTT 515_a12.seq
5197 TGCTGTTTTACCTTCTTGGCCTTTTTGTTTTTCTTTAGTCAAGTCGTTAT cjb111_a12.seq
5194 AGCTGTTTTACCTTTAGCACCTTTTTGATCGTCTGCTGAAAAGTTTGTTGT h36b_a12.seq

     AAGCTGCGATAGCTTCATCTAAAGCTATTTTAGTAGTAGCTAAAGCTGTT Majority
              5260      5270      5280      5290      5300
5238 AAGCAGCGATAGCTTCATCTAAAGCTATTTTCTTAGCAGCTAAAG---TT 2603_a12.seq
5238 AAGCAGCGATAGCTTCATCTAAAGCTATTTTCTTAGCAGCTAAAG---TT 18rs21_a12.seq
5248 AAGCTTTAACAGCAGCATCCAAAGCTGATTTAGTTGAACTACAGCAGCT 515_a12.seq
5247 AAGCTTTAACAGCTTCATCCAATGCTAGTTTAGCAGTTTTACAGCTGCC cjb111_a12.seq
5244 AAGCTGCGATTTTAGCGTCAAAGAGCAGTTTTAGCATTTGTTAGAAGCTGTT h36b_a12.seq

     TTTTGAGCTGCTTTTGCTTCATCTGCTTTAAGTGCAAGGTATTTACCTGC Majority
              5310      5320      5330      5340      5350
5285 TTTTGACCT---TCTGATTGATCTGCTTTAAGAGCAAGGTATTTACCTGC 2603_a12.seq
5285 TTTTGACCT---TCTGATTGATCTGCTTTAAGAGCAAGGTATTTACCTGC 18rs21_a12.seq
5298 TTTTCTGCATCTGTTGCAACAGCTGATTTAGGTGCCAAGGTACTTGCCATC 515_a12.seq
5297 TTTGCTTCAGCAGTTGCTGCAGCTGCTTTAGGTGCCAAGGTATTTGCCTTC cjb111_a12.seq
5294 GTTTGAGCTGATATTGTCGCAGATGATTTCAATGCAAGGTATTTACCAGC h36b_a12.seq

     TGAGTTTTTAACAACGAATTGTGCACCTGCCAAGCGTTCAGTTTCAGGTT Majority
              5360      5370      5380      5390      5400
5332 TGAGTTTTTCACAACGAATTGTGCACCAGCCAAACGGTCACCTT---GTT 2603_a12.seq
5332 TGAGTTTTTCACAACGAATTGTGCACCAGCCAAACGGTCACCTT---GTT 18rs21_a12.seq
5348 T---TTCTTAACAAGGAAGGTAGCCCCTGCCAAGCGTTCTTTTCCAAGTT 515_a12.seq
5347 T---TTCTTAACGAGGAAGGTAGCCCCTGCCAAGCGTTCAGTGTTAGCTT cjb111_a12.seq
5344 TGAATGTTTAACAACGAAGTGTGCACCTTGCAAGCGTTCTGTTTCAGTCG h36b_a12.seq
```

FIGURE 19K

	5410 · 5420 · 5430 · 5440 · 5450	
	CATTTGTTTTGACAAATTTCTTACCATGAGTCACAACTTTTGGTTCAGTT	Majority
5379	CATTAGTTTTGACAAATTTCTTACCATGAGTTTCAACTTTTGGTTCAGTT	2603_ai2.seq
5379	CATTAGTTTTGACAAATTTCTTACCATGAGTTTCAACTTTTGGTTCAGTT	18rs21_ai2.seq
5395	TATTTGTTTTCACAAATTTACGTCCATAAGTCACCACTTTTGGTTCTGAT	515_ai2.seq
5394	GATTTGTTTTCACAAATTTACGTCCATAAGTCACCACTTTTGGTTCTGAT	cjb111_ai2.seq
5394	AACTTGTTTTGACGAATTTTTTACCGTGTGTCACAACTTTTGGTTCAGTT	h36b_ai2.seq

	5460 · 5470 · 5480 · 5490 · 5500	
	GGGTTGATTGGTGTTGGGTTATCTGAATCTTTGGTATTGGTGATGGTTAC	Majority
5429	GGGTTCAATGGTGTTGGGTTATCAGAATCTTTGGTATTGGTAATGGTTAC	2603_ai2.seq
5429	GGGTTCAATGGTGTTGGGTTATCAGAATCTTTGGTATTGGTAATGGTTAC	18rs21_ai2.seq
5445	GGGTTGATTGGAGTTGGCTCATTTGAGTCTTTGTTGTTCTTGATGGTTAC	515_ai2.seq
5444	GGGTTGATTGGAGTTGGATCATTTGAGTTTTTGTTGTTCTTGATAGTCAC	cjb111_ai2.seq
5444	GGATTAAGTGGTGTTGGGTTGTCTGAATCTGGAGTATTGGTGATAGCTGC	h36b_ai2.seq

	5510 · 5520 · 5530 · 5540 · 5550	
	ATTACCATTTTCAAGAGTTAT---TGCAGTACCGTAACCAGTAACACGTT	Majority
5479	TTTACCATTTTCTAGATTTAT---TGCACTTCCGTAACCAGAAACACGTT	2603_ai2.seq
5479	TTTACCATTTTCTAGATTTAT---TGCACTTCCGTAACCAGAAACACGTT	18rs21_ai2.seq
5495	AACGCCATTTACAAATGAGACGTATTCTGGAGCGTAGCCGCTAACACGTT	515_ai2.seq
5494	AACACCATTTTTAAATGATACGTATTCTGGAGTGTAGCCGCTAACACGTT	cjb111_ai2.seq
5494	ACTACCAGCAGCAGTAGTGAT---TGTATTACCATAACCAGCGATACGTT	h36b_ai2.seq

	5560 · 5570 · 5580 · 5590 · 5600	
	CTGAGATCATGTATGTTTTATTATCATCCAGACCAGTGAATTTACCTGCG	Majority
5526	CTGAGATCATGTATGATTTGTTTTC---TAGACCAGTGAATTTACCCGAG	2603_ai2.seq
5526	CTGAGATCATGTATGATTTGTTTTC---TAGACCAGTGAATTTACCCGAG	18rs21_ai2.seq
5545	CGATAACGCGGTAAGTTTTAGCATTATCCAAGTTTTCAAAAGTGTGTTTG	515_ai2.seq
5544	CGACAACGCGGTAAGTTTTAGCATTATCCAAACCTGTGAAAGTATGTTCA	cjb111_ai2.seq
5541	CTGAGATCATGTATGTTTTACTATCAGTCAGACCACTGAATTTACCAGCA	h36b_ai2.seq

	5610 · 5620 · 5630 · 5640 · 5650	
	AAGTTACCAGATACTGTAAATTTGATACCATTACCAATGTCGATTGTACC	Majority
5573	AAGTTACCAGATACTTCAAATTTGATACCATTTCCAAGGTCGATTGTACC	2603_ai2;seq
5573	AAGTTACCAGATACTTCAAATTTGATACCATTTCCAAGGTCGATTGTACC	18rs21_ai2.seq
5595	AAGCTAGTTGCAGCTGTTGCTTTAGCTGAATCAACATTCACCCATTTGTC	515_ai2.seq
5594	AAGCGTGATGGTTTTGTTGCTTCGTGTGAAGCAACGTTCACCCATGTACC	cjb111_ai2.seq
5591	AACGCTCCAGTAACTGTAAATTTAATACCATTACCTAAGTTGATTTCGCC	h36b_ai2.seq

	5660 · 5670 · 5680 · 5690 · 5700	
	ATT---AGGT---GTTTTTGTCAATGATACTGAAGCAACAGCTGTATCAT	Majority
5623	TTT---AGAT---GTTTTTGTCAATGATACTGAAGCAACAGTTTTATCTT	2603_ai2.seq
5623	TTT---AGAT---GTTTTTGTCAATGATACTGAAGCAACAGTTTTATCTT	18rs21_ai2.seq
5645	ACC---ATCT---TTAACTTGCAAGGTGAAGATAGCATCAACTGTTTCAT	515_ai2.seq
5644	ATCCGTTTGT---TTTTCTTGCAAGGTAAAGATAGCTTTAACTGCAACAT	cjb111_ai2.seq
5641	ATT---TGGTGTTGTTTTTGTCAATGAAACTGAGGCAACAGCTGTTCCAC	h36b_ai2.seq

	5710 · 5720 · 5730 · 5740 · 5750	
	TATCTTTAAATGTGTAAACAACGTTTACATTATCTGGTTCACTACCTTCT	Majority
5667	TATCTTTCAATGTGTAAACAACGTTTACACCATCAGGTGCAATTCCGTCA	2603_ai2.seq
5667	TATCTTTCAATGTGTAAACAACGTTTACACCATCAGGTGCAATTCCGTCA	18rs21_ai2.seq
5689	CTGCTTTA---------------------TTCACTTCATTGCCATCTACT	515_ai2.seq
5691	TAGCATCA---------------------GTAATTG---TACCATCTACT	cjb111_ai2.seq
5688	TATCTTTAAGGGTATAAACAACATTCGCATTTTCTAAATCTGAACCTTTG	h36b_ai2.seq

	5760 · 5770 · 5780 · 5790 · 5800	
	GCCCAAGTTTTAGTAACTGTTATTTCACCCTTTGATGGTGTAACTGGTAG	Majority
5717	GACCAAGTTTTAGCAACTGTTACTTCACCCTTTGAAGGTGTAACAGGAAG	2603_ai2.seq
5717	GACCAAGTTTTAGCAACTGTTACTTCACCCTTTGAAGGTGTAACAGGAAG	18rs21_ai2.seq
5718	GCCCATGTTTTGTTAACAGTGATTTTCTTATCAACTGGAATACCTTCTTT	515_ai2.seq
5717	GCCCAGTCTTTAATGACTTTAATTTCTTGGTTAGCTGGAGTACCTTCTTG	cjb111_ai2.seq
5738	TCCCAAGTTTTGCTAACAGTGATTTCGCCATTTGACGGTGTTACTGGGAT	h36b_ai2.seq

	5810 · 5820 · 5830 · 5840 · 5850	
	TTCTTTCATTTCTTTACCTGGTTTGTTACCATAGTCCAATTTGATATCAT	Majority
5767	TTCAGTCAAGTCTTTACCTGGTTTGTTACCATACGACAATTTGATATCAT	2603_ai2.seq
5767	TTCAGTCAAGTCTTTACCTGGTTTGTTACCATACGACAATTTGATATCAT	18rs21_ai2.seq
5768	TGGTTCATTTTCAATTGTTGGGTTGTTGCCTAGTCCAATTTAACATCAT	515_ai2.seq
5767	TGGTTCACTTTCTTCCGTTGGGTTATTACCATAGTCCAATTTAACATCAT	cjb111_ai2.seq
5788	TTCTTTAACTTTTTTACCTGGTTTGTTACCGTATTGAAGTTTGATATCAT	h36b_ai2.seq

FIGURE 19L

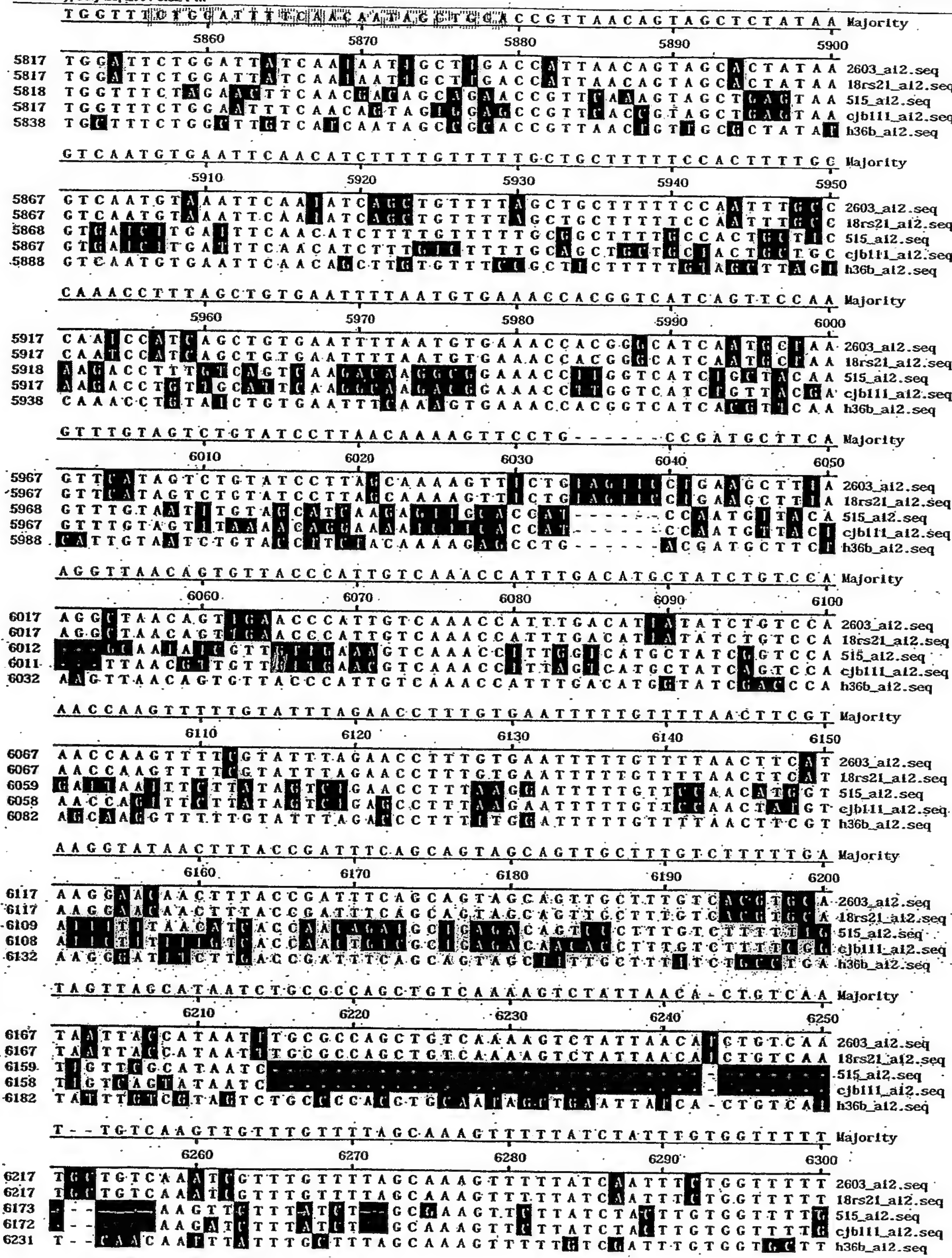


FIGURE 19M

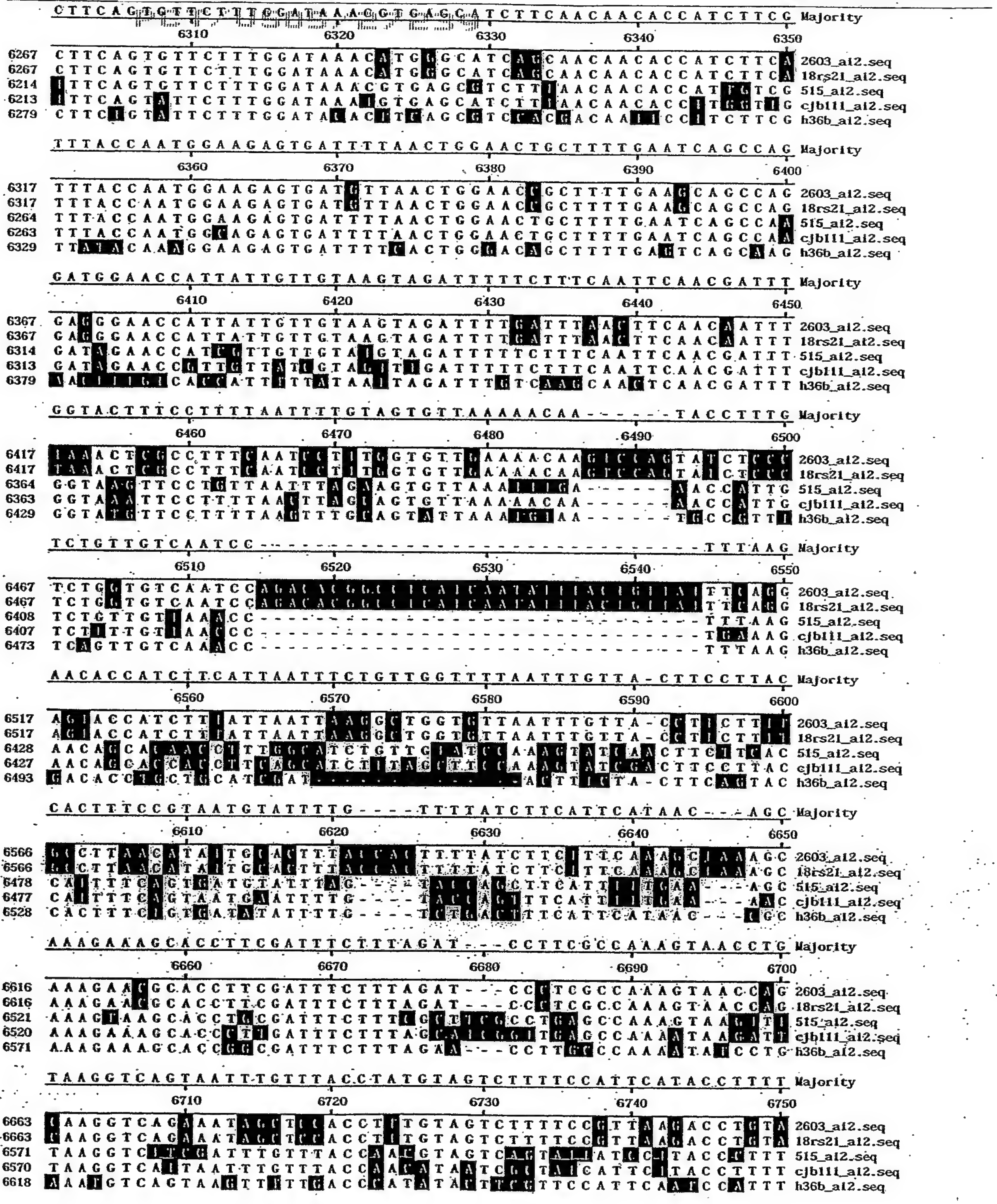


FIGURE 19N

```
     GTTCCTGGGAAGTTACTTTTGTCAAATTAGCTTGTGTTTGGACAATCTT Majority
               6760      6770      6780      6790      6800
6713 GTTCCTGGGAAGTTACTTTTGTTAAGATTTGATTCGGTTTGCAAAATCTT 2603_a12.seq
6713 GTTCCTGGGAAGTTACTTTTGTTAAGATTTGATTCGGTTTGCAAAATCTT 18rs21_a12.seq
6621 GTACCAGCAG---TAAAACTGTCAAATGCAGTTCGAGGCATGACAATCTT 515_a12.seq
6620 GTACCTTCAG---TAAAGTTATCAAATGCAGCTTGTGGCATGACAATCTT cjb111_a12.seq
6668 ATTCCTGGGAAATTTGGCTTAGCCAAATTATCGGATGTTTGAACAATTTT h36b_a12.seq

     GTGCAAGGTCACTGTATTAGT---TGT------TGCTTCATCCGCAAACG Majority
               6810      6820      6830      6840      6850
6763 GTGCAAAGTCACTGTATTAGT---TGT------TGCTTCATCCGCAAACG 2603_a12.seq
6763 GTGCAAAGTCACTGTATTAGT---TGT------TGCTTCATCCGCAAACG 18rs21_a12.seq
6668 GTGCAAGGTCACTGTGTCAGTAGTTTT------TGCTTCTTCCGCAAACA 515_a12.seq
6667 GTGCAAGGTCACAGTATCAGT---TGT------TGCGTCGTCCGCAAACG cjb111_a12.seq
6718 GTGTAAGGTTACTGTTTTAGT---GATATTTCCATCTCAGCAGCAAATA  h36b_a12.seq

     CTGGTGCAACTGAGAACAATGACGTTAAAGTCAGTAACAATGCCGAGAAC Majority
               6860      6870      6880      6890      6900
6804 CTGGTGCAACTGAGAGCAATGACGTTAAAGTCAGTAACAATGCCGAGAAC 2603_a12.seq
6804 CTGGTGCAACTGAGAGCAATGACGTTAAAGTCAGTAACAATGCCGAGAAC 18rs21_a12.seq
6712 CTGGTGCAACTGAGAACAATGATGTTAGGGTCAGTAGCAATGCCGAGAAG 515_a12.seq
6708 CTGGTGCAACTGAGAATATGACGTTAAGATCAATAGCAGTGTCGAGAAC  cjb111_a12.seq
6765 CTGGAGCTACTGATAACAAGGATGTTAAAATTAATAACAATGCCGAGAAC h36b_a12.seq

     ATTGCAAAATATTTGTTGATTCTTTTCATTTCTATCTCCTTCTTATTTTA Majority
               6910      6920      6930      6940      6950
6854 ATTGCAAAATATTTGTTGATTCTTTTCATTTCTATCTCCTTCTTATTTTA 2603_a12.seq
6854 ATTGCAAAATATTTGTTGATTCTTTTCATTTCTATCTCCTTCTTATTTTA 18rs21_a12.seq
6762 ACTGCAAAATATTTGTTGATTTTTTTCATTTCTATCTCCTTCTTATTTTA 515_a12.seq
6758 ATTGTAAGAAATTTGTTGATTTTTTTCATTTCTATCTCCTTCTTATTCTA cjb111_a12.seq
6815 ATTGCAAAATATTTGTTGATTCTTTTCATTTCTATCTCCTTCTTATTTTA h36b_a12.seq

     GTTAATCAACATGATTAATAATATGCGGATTTTAATA-C-ACCGCAGCAC Majority
               6960      6970      6980      6990      7000
6904 GTTAATCAACATGATTAATAATATGCGGATTTTAATA-C-ACCGCAGCAC 2603_a12.seq
6904 GTTAATCAACATGATTAATAATATGCGGATTTTAATA-C-ACCGCAGCAC 18rs21_a12.seq
6812 GTTAATCAACATGATTAATAATATGCGGATTTTAATA-C-ACCGCAGCAC 515_a12.seq
6808 GTTAATCAACATAGATAATAATATACGGATTATAATATT-ACCGCAGCAC cjb111_a12.seq
6865 GTTAATCAACATAAATAATAATATACGGATTATAATA-CTACCGCAGCAC h36b_a12.seq

     CACTCCCTTCAAGTCATGGAATTTTAGTTAATTAATTAAGAATACTAAAG Majority
               7010      7020      7030      7040      7050
6952 CACTCCCTTCAAGTCATGGAATTTTAGTTAATTAATTAAGAATACTAAAG 2603_a12.seq
6952 CACTCCCTTCAAGTCATGGAATTTTAGTTAATTAATTAAGAATACTAAAG 18rs21_a12.seq
6860 CACTCCCTTCAAGTCATGGAATTTTAGTTAATTAATTAAGAATACTAAAG 515_a12.seq
6857 CACTCCTTTC-------GAATTATATTTAATTAATTAAGAATACTAAAA  cjb111_a12.seq
6914 CACTCCTTTCAAAAT-------------ATTTAATTAAAAATGCTAAAG  h36b_a12.seq

     CGCATAATTTTTAATCTTTTTTGATGGACATATCACTAGATTTCTTATAC Majority
               7060      7070      7080      7090      7100
7002 CGCATAATTTTTAATCTTTTTTGATGGACATATCACTAGATTTCTTATAC 2603_a12.seq
7002 CGCATAATTTTTAATCTTTTTTGATGGACATATCACTAGATTTCTTATAC 18rs21_a12.seq
6910 CGCATAATTTTTAATCTTTTTTGATGGACATATCACTAGATTTCTTATAC 515_a12.seq
6899 CACATGATTCTTAATCTTTTTCTCTAGAATATATCACTAGATTTCTTATAT cjb111_a12.seq
6952 CACATAGTTCTTAATCTTTCTTGATTGAGCATCACTAGATTTCTTATGT  h36b_a12.seq

     CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGCTCCACCTAT Majority
               7110      7120      7130      7140      7150
7052 CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGCTCCACCTAT 2603_a12.seq
7052 CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGCTCCACCTAT 18rs21_a12.seq
6960 CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGCTCCACCTAT 515_a12.seq
6949 CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGATCCACCTAT cjb111_a12.seq
7002 CTTTTCCAAATATAAATTCCACCTGCAATAGACATCATAGCTCCACCTAT h36b_a12.seq

     TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCATAGGAATAATTCCTT Majority
               7160      7170      7180      7190      7200
7102 TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCATAGGAATAATTCCTT 2603_a12.seq
7102 TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCATAGGAATAATTCCTT 18rs21_a12.seq
7010 TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCGTAGGAATAATTCCTT 515_a12.seq
6999 TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCATCGGAATAATTCCTT cjb111_a12.seq
7052 TAAAATGAAAGATAGAATTCCTTTCCCACCTGTCATCGGAATAATTCCTT h36b_a12.seq
```

FIGURE 19O

```
     TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA Majority
          7210      7220      7230      7240      7250
7152 TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA 2603_a12.seq
7152 TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA 18rs21_a12.seq
7060 TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA 515_a12.seq
7049 TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA cjb111_a12.seq
7102 TTGGTGGAATATGCGTGTTGGTAATTAAATGCTTGTCACCTTCCTCATGA h36b_a12.seq

     TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT Majority
          7260      7270      7280      7290      7300
7202 TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 2603_a12.seq
7202 TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 18rs21_a12.seq
7110 TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTTTATCGATCCTTT 515_a12.seq
7099 TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTGTATCGATCCTTT cjb111_a12.seq
7152 TATTCAGAAATCTGTTTATTAACAGCTATTATATTTTGTATCGATCCTTT h36b_a12.seq

     AACCACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT Majority
          7310      7320      7330      7340      7350
7252 AACCACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT 2603_a12.seq
7252 AACCACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT 18rs21_a12.seq
7160 AACCACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT 515_a12.seq
7149 AACAACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT cjb111_a12.seq
7202 AACAACTTCAAAAGTTAAAATTGGTTTATTAGTAATTTTTTGATAATCCT h36b_a12.seq

     CCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG Majority
          7360      7370      7380      7390      7400
7302 CCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG 2603_a12.seq
7302 CCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG 18rs21_a12.seq
7210 CCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG 515_a12.seq
7199 TCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG cjb111_a12.seq
7252 TCGGCGAAACTGCTTCTATTAACTGATATTTGCCATCTTTCAAATCTTTG h36b_a12.seq

     TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT Majority
          7410      7420      7430      7440      7450
7352 TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT 2603_a12.seq
7352 TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT 18rs21_a12.seq
7260 TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT 515_a12.seq
7249 TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT cjb111_a12.seq
7302 TAAGAAATTTTGCCGTTTTCTCCCGTCACTACTTTTGAATTATTATTTTT h36b_a12.seq

     TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAACG Majority
          7460      7470      7480      7490      7500
7402 TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAACG 2603_a12.seq
7402 TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAACG 18rs21_a12.seq
7310 TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAACG 515_a12.seq
7299 TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAACG cjb111_a12.seq
7352 TATTGGTAAATAAAGTTTATAATCTTCATTAAATTCTTGAAGTTCAAATG h36b_a12.seq

     TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT Majority
          7510      7520      7530      7540      7550
7452 TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTATAAAT 2603_a12.seq
7452 TAGCTCCTTTGAGAAGCAATTTATTATTATCTTTATCAACTTTTATAAAT 18rs21_a12.seq
7360 TAGCTCCTTTGAGGAGCAACTTATTATTATCTTTATCAACTTTTATAAAT 515_a12.seq
7349 TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTGTAAAT cjb111_a12.seq
7402 TAGCTCCTTTGAGAAGCAACTTATTATTATCTTTATCAACTTTTGTAAAT h36b_a12.seq

     TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC Majority
          7560      7570      7580      7590      7600
7502 TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC 2603_a12.seq
7502 TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC 18rs21_a12.seq
7410 TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC 515_a12.seq
7399 TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC cjb111_a12.seq
7452 TCAATTTCACCTAACTTCTTCTCGTTTTTAATCGTTATTGTAGGATATTC h36b_a12.seq

     TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATTAGGAT Majority
          7610      7620      7630      7640      7650
7552 TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATTAGGAT 2603_a12.seq
7552 TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATTAGGAT 18rs21_a12.seq
7460 TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATTAGGAT 515_a12.seq
7449 TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATTAGGAT cjb111_a12.seq
7502 TCTCACATCACGAATTTTAGGGATTGGAAAATCTCTAAGTGTATCGGGTT h36b_a12.seq
```

FIGURE 19P

```
     CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG Majority
           7660      7670      7680      7690      7700
7602 CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG 2603_a12.seq
7602 CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG 18rs21_a12.seq
7510 CCTCTGACTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG 515_a12.seq
7499 CCTCTGATTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG cjb111_a12.seq
7552 CCTCTGATTTAGGATTCAATGTTGTTCTACCATTAGTGTCATAGAATTTG h36b_a12.seq

     TTACTTATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT Majority
           7710      7720      7730      7740      7750
7652 TTACTTATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT 2603_a12.seq
7652 TTACTTATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT 18rs21_a12.seq
7560 TTACTTATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT 515_a12.seq
7549 TTACTTATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT cjb111_a12.seq
7602 TTACTAATAAAACTGTCATCTAGTTTCACATCATATGTGAGTGTTACTTT h36b_a12.seq

     TTGACCTTCTCCTAAATTCAAACCTCTAACATAGAGTTTATTTCCGATGT Majority
           7760      7770      7780      7790      7800
7702 TTGACCTTCTCCTAAATTCAAACCTCTAACATAGAGTTTATTTCCGATGT 2603_a12.seq
7702 TTGACCTTCTCCTAAATTCAAACCTCTAACATAGAGTTTATTTCCGATGT 18rs21_a12.seq
7610 TTGACCTTCTCCTAAATTCAAACCTCTAACATAGAGTTTATTTCCGATGT 515_a12.seq
7599 TTGTCCCTCCCCTAAGTTCAAACCTCTAACGTAGAGTTTATTTTTGATGT cjb111_a12.seq
7652 TTGTCCCTCCCCTAAGTTCAAACCTCTAACGTAGAGTTTATTTTTGATGT h36b_a12.seq

     ATTCTAATTTAACCCCCTTAAGTATTCCACCATCATTATTAGGCCCACCA Majority
           7810      7820      7830      7840      7850
7752 ATTCTAATTTAACCCCCTTAAGTATTCCACCATCATTATTAGGCCCACCA 2603_a12.seq
7752 ATTCTAATTTAACCCCCTTAAGTATTCCACCATCATTATTGGGCCCACCA 18rs21_a12.seq
7660 ATTCTAATTTAACCCCCTTAAGTATTCCACCATCATTATTAGGCCCACCA 515_a12.seq
7649 ATTCTAATTTAACCCCTTTAAGTATTCCACCATCATTATTAGGCCCACCA cjb111_a12.seq
7702 ATTCTAATTTAACCCCTTTAAGTATCCCACCATCATTATTAGGCCCGCCA h36b_a12.seq

     GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCCTGTAAAGTATA Majority
           7860      7870      7880      7890      7900
7802 GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCCTGTAAAGTATA 2603_a12.seq
7802 GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCCTGTAAAGTATA 18rs21_a12.seq
7710 GTTGCAATACCATCCTTCATTACACTTCCATCATTTCCCTGTAAAGTATA 515_a12.seq
7699 GTTGCAATGCTATCTTTCATTATACTTCCATCATTTCCCTGTAAAGTATA cjb111_a12.seq
7752 GTTGCAATGCTATCTTTCATTATACTTCCATCATTTCCCTGTAAAGTATA h36b_a12.seq

     ATCACTTGGCTGTAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT Majority
           7910      7920      7930      7940      7950
7852 ATCACTTGGCTGTAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT 2603_a12.seq
7852 ATCACTTGGCTGTAATGTTTGTCCATTACCAAGCTGTAAATTGATTTTAT 18rs21_a12.seq
7760 ATCACTTGGCTGTAATATTTGTCCATTACCAAGCTGTAAATTGATTTTAT 515_a12.seq
7749 ATCACTTGGTTGCAATGTTTGTCCGTTGCCAAGCTGTAAATTGATTTTGT cjb111_a12.seq
7802 ATCACTTGGTTGCAATGTTTGTCCGTTGCCAAGATGTAAATTGATTTTAT h36b_a12.seq

     CACCCATAGGATCTTCGATAGTTCCATTAACAATTGAGTTTTCTTTTGTT Majority
           7960      7970      7980      7990      8000
7902 CACCCATAGGATCTTCGATAGTTCCATTAACAATTGAGTTTTCTTTTGTT 2603_a12.seq
7902 CACCCATAGGATCTTCGATAGTTCCATTAACAATTGAGTTTTCTTTTGTT 18rs21_a12.seq
7810 CACCCATAGGATCTTCGATAGTTCCATTAACAATTGAGTTTTCTTTTGTT 515_a12.seq
7799 CACCCATAGGATCTTCAATAGTTCCATTAACAATTGAGTTTTCTTTTGTT cjb111_a12.seq
7852 CACCCATAGGATCTTCAATAGTTCCATTAACAATTGAGTTTTCCTTTGTT h36b_a12.seq

     AAAATCGTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTGTTAGA Majority
           8010      8020      8030      8040      8050
7952 AAAATCGTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTGTTAGA 2603_a12.seq
7952 AAAATCGTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTGTTAGA 18rs21_a12.seq
7860 AAAATCGTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTGTTAGA 515_a12.seq
7849 AAAATCTTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTGTTAGA cjb111_a12.seq
7902 AAAATCTTTTCAAATTGTTGCTGAATTTTAGATAAAATTTCATTATTAGA h36b_a12.seq

     TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGGTTTGGAAGAGA Majority
           8060      8070      8080      8090      8100
8002 TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGGTTTGGAAGAGA 2603_a12.seq
8002 TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGGTTTGGAAGAGA 18rs21_a12.seq
7910 TGTATCGGCTGAAGTTACGATAGGGGTGTAGTACTCAGGTTTGGAAGAGA 515_a12.seq
7899 TGCATCGGATGAAGTTACTATCGGGGTATAATACTCAGGTTTAGAAGAGA cjb111_a12.seq
7952 TACATCAGCTGAAGTTACTATCGGGGTATAATACTCAGGTTTAGAAGAGA h36b_a12.seq
```

FIGURE 19Q

```
     A C G A C C T C A T T A G T T C T G T G A T T T C T C C A T C T G A A A G T T T A A A A G C T T C C Majority
               8110                8120                8130                8140                8150
8052 A C G A C C T C A T T A G T T C T G T G A T T T C T C C A T C T G A A A G T T T A A A A G C T T C C 2603_al2.seq
8052 A C G A C C T C A T T A G T T C T G T G A T T T C T C C A T C T G A A A G T T T A A A A G C T T C C 18rs21_al2.seq
7960 A C G A C C T C A T T A G T T C T G T G A T T T C T C C A T C T G A A A G T T T A A A A G C T T C C 515_al2.seq
7949 A T G A C T T C A T T A G T T C T G T T A T T T C C C C A T C T G A A A G T T C A A A A G C T T C C cjb111_al2.seq
8002 A T G A A T T C A T T A G T T C T G T T A T T T C C C C A T C T G A A A G T T C A A A A G C T T C C h36b_al2.seq

     T C T T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A C T C C T C A T T Majority
               8160                8170                8180                8190                8200
8102 T C T T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A C T C C T C A T T 2603_al2.seq
8102 T C T T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A C T C C T C A T T 18rs21_al2.seq
8010 T C T T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A C T C C T C A T T 515_al2.seq
7999 T C T T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A A T C C T C A T T cjb111_al2.seq
8052 T C C T T C A A T T T T T G A A A A G T A C C A T C T T G A T T T T T C T T A T A A T C C T C A T T h36b_al2.seq

     A T A A A C T T G T C T A A A A C C A G A T A T A T C G A T A C C A A A A T T A A A A A T G T C A T Majority
               8210                8220                8230                8240                8250
8152 A T A A A C T T G T C T A A A A C C A G A T A T A T C G A T A C C A A A A T T A A A A A T G T C A T 2603_al2.seq
8152 A T A A A C T T G T C T A A A A C C A G A T A T A T C G A T A C C A A A A T T A A A A A T G T C A T 18rs21_al2.seq
8060 A T A A A C T T G T C T A A A A C C A G A T A T A T C G A T A C C A A A A T T A A A A A T G T C A T 515_al2.seq
8049 A T A A A C T T G T C T A A A A C C A G A T A T A T C T A T A C C A A A A T T A A A G A T G T C A T cjb111_al2.seq
8102 A T A A A C T T G T C T A A A A C C A G A T A T A T C T A T A C C A A A A T T A A A G A T G T C A T h36b_al2.seq

     A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T Majority
               8260                8270                8280                8290                8300
8202 A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T 2603_al2.seq
8202 A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T 18rs21_al2.seq
8110 A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T 515_al2.seq
8099 A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T cjb111_al2.seq
8152 A A T T T T T C T G T T T T A A A C T A T T T A T A T A A A G T T T G G T T G G T G T T C C A T G T h36b_al2.seq

     T C T T T C A C T G G T C C A T T T C G A T A A A T T G T A C C T T T A G G G T A A T T A A G A T T Majority
               8310                8320                8330                8340                8350
8252 T C T T T C A C T G G T C C A T T T C G A T A A A T T G T A C C T T T A G G G T A A T T A A G A T T 2603_al2.seq
8252 T C T T T C A C T G G T C C A T T T C G A T A A A T T G T A C C T T T A G G G T A A T T A A G A T T 18rs21_al2.seq
8160 T C T T T C A C T G G T C C A T T T C G A T A A A T T G T A C C T T T A G G G T A A T T A A G A T T 515_al2.seq
8149 T C T C T T A C T G G T C C A T T T C G A T A A A A T G T A C C T T T A G G G T A A T T A A G A T T cjb111_al2.seq
8202 T C T C T T A C T G G T C C A T T T C T A T A A A T T G T A C C T T T A G G G T A A T T A A G A T T h36b_al2.seq

     T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A G A T T A T C T G T G T T T Majority
               8360                8370                8380                8390                8400
8302 T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A G A T T A T C T G T G T T T 2603_al2.seq
8302 T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A G A T T A T C T G T G T T T 18rs21_al2.seq
8210 T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A G A T T A T C T G T G T T T 515_al2.seq
8199 T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A G A T T A T C T G T G T T T cjb111_al2.seq
8252 T A A A T C T A A A T A A T G A A G T T T T T G T A A G T T T C C A G A A A T T A T C T G T G T T T h36b_al2.seq

     G A T A A C T A T C T A A G G G A A A C A A A A A G T A A C T C T C C C C A T T T C C T T T T A T A Majority
               8410                8420                8430                8440                8450
8352 G A T A A C T A T C T A A G G G A A A C A A A A A G T A A C T C T C C C C A T T T C C T T T T A T A 2603_al2.seq
8352 G A T A A C T A T C T A A G G G A A A C A A A A A G T A A C T C T C C C C A T T T C C T T T T A T A 18rs21_al2.seq
8260 G A T A A C T A T C T A A G G G A A A T A A A A A G T A A C T C T C C C C A T T T C C T T T T A T A 515_al2.seq
8249 G A T A A C T A T C T A A G G G A A A C A A A A A G T A A C T C T C C C C A T T T C C T T T T A T A cjb111_al2.seq
8302 G A T A A C T A T C T A A G G G A A A C A A A A A G T A A C T C T C C C C A T T G C C T T T G A T C h36b_al2.seq

     T C C T C G G G C T T A T C A G T A A G T A G A A A A T T A C T T T T A T T T A G A T A T C C A T T Majority
               8460                8470                8480                8490                8500
8402 T C C T C G G G C T T A T C A G T A A G T A G A A A A T T A C T T T T A T T T A G A T A T C C A T T 2603_al2.seq
8402 T C C T C G G G C T T A T C A G T A A G T A G A A A A T T A C T T T T A T T T A G A T A T C C A T T 18rs21_al2.seq
8310 T C A T C G G G C T T A T C A G T A A G T A G A A A A T T A C T T T T A T T T A G A T A T C C A T T 515_al2.seq
8299 T C C T C G G G C T T A T C A G T A A G T A G A A A A T T A C T T T T A T T T A G A T A T C C A T T cjb111_al2.seq
8352 T T T T C T G G A T C A T C A G T T A T A A A A T A A T T A T T T T T G T C T A A A T A A C C T T T h36b_al2.seq

     T T T T T T C A T T T G T T C A A A T T G G C T T T C A T A T G A T G C A C C C A G T T T A A A A T Majority
               8510                8520                8530                8540                8550
8452 T T T T T T C A T T T G T T C A A A T T G G C T T T C A T A T G A T G C A C C C A G T T T A A A A T 2603_al2.seq
8452 T T T T T T C A T T T G T T C A A A T T G G C T T T C A T A T G A T G C A C C C A G T T T A A A A T 18rs21_al2.seq
8360 T T T T T T C A T T T G T T C A A A T T G G C T T T C A T A T G A T G C A C C C A G T T T A A A A T 515_al2.seq
8349 T T T T T T C A T T T G T T C A A A T T G G C T T T C A T A T G A T G C A C C C A G T T T A A A A T cjb111_al2.seq
8402 T T C T T T T A T T C T C T C A A A T T G A T T T G C G T A T G T T G A A C C T T T T A C A A A A C h36b_al2.seq
```

FIGURE 19R

```
     TATTAATAGCATATGATCTCGTAGGAACACCATCAGTTACATGAACAATA Majority
              8560      8570      8580      8590      8600
8502 TATTAATAGCATATGATCTCGTAGGAACACCATCAGTTACATGAACAATA 2603_ai2.seq
8502 TATTAATAGCATATGATCTCGTAGGAACACCATCAGTTACATGAACAATA 18rs21_ai2.seq
8410 TATTAATAGCATATGATCTCGTAGGAACACCATCAGTTACATGAACAATA 515_ai2.seq
8399 TATTAATAGCATATGATCTCGTAGGAACACCATCAGTTACATGAACAATA cjb111_ai2.seq
8452 TATTAATGGCATATGATCTTGTTGGAACACCGTCAGTTAGATGAACAATA h36b_ai2.seq

     ATTTTTTGACTATTTCGATTTACTTGACTCAAAATATCATCTGCCTCCAT Majority
              8610      8620      8630      8640      8650
8552 ATTTTTTGACTATTTCGATTTACTTGACTCAAAATATCATCTGCCTCCAT 2603_ai2.seq
8552 ATTTTTTGACTATTTCGATTTACTTGACTCAAAATATCATCTGCCTCCAT 18rs21_ai2.seq
8460 ATTTTTTGACTATTTCGATTTACTTGACTCAAAATATCATCTGCCTCCAT 515_ai2.seq
8449 ATTTTTTGACTATTTCGATTTACTTGACTCAAAATATCATCTGCCTCCAT cjb111_ai2.seq
8502 ATCTTTCTACTCTTACGCTGTATACTACTTAACAAGGTATCTGCCTCCAT h36b_ai2.seq

     GAAGGCTTTCATAGTAAATGTTTCTCCTACTTTACTAAGATAGTACTCCT Majority
              8660      8670      8680      8690      8700
8602 GAAGGCTTTCATAGTAAATGTTTCTCCTACTTTACTAAGATAGTACTCCT 2603_ai2.seq
8602 GAAGGCTTTCATAGTAAATGTTTCTCCTACTTTACTAAGATAGTACTCCT 18rs21_ai2.seq
8510 GAAGGCTTTCATAGTAAATGTTTCTCCTACTTTACTAAGATAGTACTCCT 515_ai2.seq
8499 GAAGGCTTTCATAGTAAATGTTTCTCCTACTTTACTAAGATAGTACTCCT cjb111_ai2.seq
8552 AAAAGCTTTCATTGTAAAGGTCTCACCTACTTTACTTAAATCATATTCCC h36b_ai2.seq

     TTTGTTGCTCTGGAGTTAATCCATTGGTAGTAGATCCCCACTTAGCTTTA Majority
              8710      8720      8730      8740      8750
8652 TTTGTTGCTCTGGAGTTAATCCATTGGTAGTAGATCCCCACTTAGCTTTA 2603_ai2.seq
8652 TTTGTTGCTCTGGAGTTAATCCATAGGTAGTAGATCCCCACTTAGCTTTA 18rs21_ai2.seq
8560 TTTGTTGCTCTGGAGTTAATCCATTGGTAGTAGATCCCCACTTAGCTTTA 515_ai2.seq
8549 TTTGTTGCTCTGGAGTTAATCCATTGGTAGTAGATCCCCACTTAGCTTTA cjb111_ai2.seq
8602 TCTTTTTTTCTGGAGTTAATCCTAGACTTGTCCCCCCCCACTTAGCTTCT h36b_ai2.seq

     GGAGCTTCTGTCGGAATCCTTTTTATAATCTCTTCAGCATTATTTGTTAA Majority
              8760      8770      8780      8790      8800
8702 GGAGCTTCTGTCGGAATCCTTTTTATAATCTCTTCAGCATTATTTGTTAA 2603_ai2.seq
8702 GGAGCTTCTGTCGGAATCCTTTTTATAATCTCTTCAGCATTATTTGTTAA 18rs21_ai2.seq
8610 GGAGCTTCTGTCGGAATCCTTTTTATAATCTCTTCAGCATTATTTGTTAA 515_ai2.seq
8599 GGAGCTTCTGTCGGAATCCTTTTTATAATCTCTTCAGCATTATTTGTTAA cjb111_ai2.seq
8652 GGAGCTTCTTTAGGGATCTTTTTTATAATATCAGCAGCAATATTAGTGAA h36b_ai2.seq

     TTGTTTATGACTATAATTCTCTGTCTGAATTGTGAACTTAGTTTGAAGGC Majority
              8810      8820      8830      8840      8850
8752 TTGTTTATGACTATAATTCTCTGTCTGAATTGTGAACTTAGTTTGAAGGC 2603_ai2.seq
8752 TTGTTTATGACTATAATTCTCTGTCTGAATTGTGAACTTAGTTTGAAGGC 18rs21_ai2.seq
8660 TTGTTTATGACTATAATTCTCTGTCTGAATTGTGAACTTAGTTTGAAGGC 515_ai2.seq
8649 TTGTTTATGACTATAATTCTCTGTCTGAATTGTGAACTTAGTTTGAAGGC cjb111_ai2.seq
8702 CTTTTTATAGCTATAATCATTTGTCTGAACTGTGAAACTAGTTTCAAGTC h36b_ai2.seq

     CATAATATTTATCATCTTCTTTAAATCCTTTTACGACATCTACACTCCTA Majority
              8860      8870      8880      8890      8900
8802 CATAATATTTATCATCTTCTTTAAATCCTTTTACGACATCTACACTCCTA 2603_ai2.seq
8802 CATAATATTTATCATCTTCTTTAAATCCTTTTACGACATCTACACTCCTA 18rs21_ai2.seq
8710 CATAATATTTATCATCTTCTTTAAATCCTTTTACGACATCTACACTCCTA 515_ai2.seq
8699 CATAATATTTATCATCTTCTTTAAATCCTTTTACGACATCTACACTCCTA cjb111_ai2.seq
8752 CATAATA---AGGATCCTCTTTAAAACCTTTTATAACTTTAACTGTCCTT h36b_ai2.seq

     CCATCAAAAATATCTGAACCATAGGTAACTAATGCAACCCTATTATCACT Majority
              8910      8920      8930      8940      8950
8852 CCATCAAAAATATCTGAACCATAGGTAACTAATGCAACCCTATTATCACT 2603_ai2.seq
8852 CCATCAAAAATATCTGAACCATAGGTAACTAATGCAACCCTATTATCACT 18rs21_ai2.seq
8760 CCATCAAAAATATCTGAACCATAGGTAACTAATGCAACCCTATTATCACT 515_ai2.seq
8749 CCATCAAAAATATCTGAACCATAGGTAACTAATGCAACCCTATTATCACT cjb111_ai2.seq
8799 CCATCAAAAATATCTGAACCATAAGTAACTAAAGCTGCTCGGTTTTCAAC h36b_ai2.seq

     GTTTGCTCCTAAAATATCTTTTACTGCGGTCCCAAGAGCTTCGGCAGCTT Majority
              8960      8970      8980      8990      9000
8902 GTTTGCTCCTAAAATATCTTTTACTGCGGTCCCAAGAGCTTCGGCAGCTT 2603_ai2.seq
8902 GTTTGCTCCTAAAATATCTTTTACTGCGGTCCCAAGAGCTTCGGCAGCTT 18rs21_ai2.seq
8810 GTTTGCTCCTAAAATATCTTTTACTGCGGTCCCAAGAGCTTCGGCAGCTT 515_ai2.seq
8799 GTTTGCTCCTAAAATATCTTTTACTGCGGTCCCAAGAGCTTCGGCAGCTT cjb111_ai2.seq
8849 ATTTGCTCCTAAAACATCTTTTATAATTGTTTCTACTGCTTCACCTGCCT h36b_ai2.seq
```

FIGURE 19S

76/487


```
     TCTTGGCTTTATTATGCCTTTGAAAATTTGGCCCATCGTTATTCATTGAG Majority
              9010      9020      9030      9040      9050
8952 TCTTGGCTTTATTATGCCTTTGAAAATTTGGCCCATCGTTATTCATTGAG 2603_a12.seq
8952 TCTTGGCTTTATTATGCCTTTGAAAATTTGGCCCATCGTTATTCATTGAG 18rs21_a12.seq
8860 TCTTGGCTTTATTATGCCTTTGAAAATTTGGCCCATCGTTATTCATTGAG 515_a12.seq
8849 TCTTGGCTTTATTATGCCTTTGAAAATTTGGCCCATCGTTATTCATTGAG cjb111_a12.seq
8899 TTTTTGCCTTGTTATTTTTT------------CCATTATTCTTCATAGAA h36b_a12.seq

     TTAGAATTATCGAGTACGAAGACAACATCTAACGGCTTTTGTTTGTCCAC Majority
              9060      9070      9080      9090      9100
9002 TTAGAATTATCGAGTACGAAGACAACATCTAACGGCTTTTGTTTGTCCAC 2603_a12.seq
9002 TTAGAATTATCGAGTACGAAGACAACATCTAACGGCTTTTGTTTGTCCAC 18rs21_a12.seq
8910 TTAGAATTATCGAGTACGAAGACAACATCTAACGGCTTTTGTTTGTCCAC 515_a12.seq
8899 TTAGAATTATCGAGTACGAAGACAACATCTAACGGCTTTTGTTTGTCCAC cjb111_a12.seq
8937 TTTGAATTATCAAGAACAAAAACAACATCCAGAGGTTCATCTTTATTTAT h36b_a12.seq

     TGGTTTTACTATGGTTTTCCACTGACAGTTAACTCAATTTTATATTTAT Majority
              9110      9120      9130      9140      9150
9052 TGGTTTTACTATGGTTTTCCACTGACAGTTAACTCAATTTTATATTTAT 2603_a12.seq
9052 TGGTTTTACTATGGTTTTCCACTGACAGTTAACTCAATTTTATATTTAT 18rs21_a12.seq
8960 TGGTTTTACTATGGTTTTCCACTGACAGTTAACTCAATTTTATATTTAT 515_a12.seq
8949 TGGTTTTACTATGGTTTTCCACTGACAGTTAACTCAATTTTATATTTAT cjb111_a12.seq
8987 AGTTTTTATTATGGATTTACCGCTAACAGTTAACTCAATTTTATATTTAT h36b_a12.seq

     TATGAGCTAAATCACCTACTTCTGAAATACGTTTAGATAATGTTCCCTCT Majority
              9160      9170      9180      9190      9200
9102 TATGAGCTAAATCACCTACTTCTGAAATACGTTTAGATAATGTTCCCTCT 2603_a12.seq
9102 TATGAGCTAAATCACCTACTTCTGAAATACGTTTAGATAATGTTCCCTCT 18rs21_a12.seq
9010 TATGAGCTAAATCACCTACTTCTGAAATACGTTTAGATAATGTTCCCTCT 515_a12.seq
8999 TATGAGCTAAATCACCTACTTCTGAAATACGTTTAGATAATGTTCCCTCT cjb111_a12.seq
9037 TATGATCCAAATCATTTACTTCTGAAATACGTTTAGATAATGTTCCCTCT h36b_a12.seq

     GGAATTTCTCTTATATGCTCACCTTCACTTGAATATGGGTTAACTGCTTT Majority
              9210      9220      9230      9240      9250
9152 GGAATTTCTCTTATATGCTCACCTTCACTTGAATATGGGTTAACTGCTTT 2603_a12.seq
9152 GGAATTTCTCTTATATGCTCACCTTCACTTGAATATGGGTTAACTGCTTT 18rs21_a12.seq
9060 GGAATTTCTCTTATATGCTCACCTTCACTTGAATATGGGTTAACTGCTTT 515_a12.seq
9049 GGAATTTCTCTTATATGCTCACCTTCACTTGAATATGGGTTAACTGCTTT cjb111_a12.seq
9087 TGAATTTCTCTTATGTGCTCACCTTCACTTGAATATGGATTAACTGCTTT h36b_a12.seq

     TGCCTCTGACTTTCCATTTGGAACTGAACCTTTAACATGCTCAAGTTTAT Majority
              9260      9270      9280      9290      9300
9202 TGCCTCTGACTTTCCATTTGGAACTGAACCTTTAACATGCTCAAGTTTAT 2603_a12.seq
9202 TGCCTCTGACTTTCCATTTGGAACTGAACCTTTAACATGCTCAAGTTTAT 18rs21_a12.seq
9110 TGCCTCTGACTTTCCATTTGGAACTGAACCTTTAACATGCTCAAGTTTAT 515_a12.seq
9099 TGCCTCTGACTTTCCATTTGGAACTGAACCTTTAACATGCTCAAGTTTAT cjb111_a12.seq
9137 TGCCTCTAATTTCCCATTTGGAATTGAATTTTTAACATGCTCAAGATTAT h36b_a12.seq

     AAGATTCCTTTGTATCTTCATAAATTCCTGTGGGGGATACTGCTTATCT Majority
              9310      9320      9330      9340      9350
9252 AAGATTCCTTTGTATCTTCATAAATTCCTGTGGGGGATACTGCTTATCT 2603_a12.seq
9252 AAGATTCCTTTGTATCTTCATAAATTCCTGTGGGGGATACTGCTTATCT 18rs21_a12.seq
9160 AAGATTCCTTTGTATCTTCATAAATTCCTGTGGGGGATACTGCTTATCT 515_a12.seq
9149 AAGATTCCTTTGTATCTTCATAAATTCCTGTGGGGGATACTGCTTATCT cjb111_a12.seq
9187 AAGATTCTTTTGTATCTTCATAAGCTCCTGTAAGGGATACTGCTTATCT h36b_a12.seq

     AGTTCTTCCTGATTTTGTCCAATTGTGGAATTTTTATCACCACTATTTTG Majority
              9360      9370      9380      9390      9400
9302 AGTTCTTCCTGATTTTGTCCAATTGTGGAATTTTTATCACCACTATTTTG 2603_a12.seq
9302 AGTTCTTCCTGATTTTGTCCAATTGTGGAATTTTTATCACCACTATTTTG 18rs21_a12.seq
9210 AGTTCTTCCTGATTTTGTCCAATTGTGGAATTTTTATCACCACTATTTTG 515_a12.seq
9199 AGTTCTTCCTGATTTTGTCCAATTGTGGAATTTTTATCACCACTATTTTG cjb111_a12.seq
9237 AGTTCCTCTTGCCTTTGTTCAATTATAGATTTTTTATCATCACTATTTTG h36b_a12.seq

     TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGTTAG Majority
              9410      9420      9430      9440      9450
9352 TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGTTAG 2603_a12.seq
9352 TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGTTAG 18rs21_a12.seq
9260 TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGTTAG 515_a12.seq
9249 TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGTTAG cjb111_a12.seq
9287 TATCGTAGTTTTTCCATTACTCTCAACCTTAACTTGCCAAGTCTGGGTAG h36b_a12.seq
```

FIGURE 19T

```
     TCTTTTTATAACCTTCGGGCGCTGTTTCTTCTGATAAAGTATAATCTCCA Majority
              9460      9470      9480      9490      9500
9402 TCTTTTTATAACCTTCGGGCGCTGTTTCTTCTGATAAAGTATAATCTCCA 2603_ai2.seq
9402 TCTTTTTATAACCTTCGGGCGCTGTTTCTTCTGATAAAGTATAATCTCCA 18rs21_ai2.seq
9310 TCTTTTTATAACCTTCGGGCGCTGTTTCTTCTGATAAAGTATAATCTCCA 515_ai2.seq
9299 TCTTTTTATAACCTTCGGGCGCTGTTTCTTCTGATAAAGTATAATCTCCA cjb111_ai2.seq
9337 TCTTTTTGTATCCTTCGGGTGCCGTTTCTTCTGATAAAGTGTAATCTCCA h36b_ai2.seq

     GGTATGAGATTATCAAAAGTAGCTTCACCTGTTAGCTCAGCAGTTACTTT Majority
              9510      9520      9530      9540      9550
9452 GGTATGAGATTATCAAAAGTAGCTTCACCTGTTAGCTCAGCAGTTACTTT 2603_ai2.seq
9452 GGTATGAGATTATCAAAAGTAGCTTCACCTGTTAGCTCAGCAGTTACTTT 18rs21_ai2.seq
9360 GGTATGAGATTATCAAAAGTAGCTTCACCTGTTAGCTCAGCAGTTACTTT 515_ai2.seq
9349 GGTATGAGATTATCAAAAGTAGCTTCACCTGTTAGCTCAGCAGTTACTTT cjb111_ai2.seq
9387 GGTGTGAGATTATCAAAAGTAGCTTCACCTGTTACCTCAGTAGTTACTTT h36b_ai2.seq

     TTCTATTTTACTTTCTGGATGAGCAGTAGTTTTTAAAACAAAGGTAGCTT Majority
              9560      9570      9580      9590      9600
9502 TTCTATTTTACTTTCTGGATGAGCAGTAGTTTTTAAAACAAAGGTAGCTT 2603_ai2.seq
9502 TTCTATTTTACTTTCTGGATGAGCAGTAGTTTTTAAAACAAAGGTAGCTT 18rs21_ai2.seq
9410 TTCTATTTTACTTTCTGGATGAGCAGTAGTTTTTAAAACAAAGGTAGCTT 515_ai2.seq
9399 TTCTATTTTACTTTCTGGATGAGCAGTAGTTTTTAAAACAAAGGTAGCTT cjb111_ai2.seq
9437 TTCTACTTTGCTTTCTGAGTGTGAAGTAGGTTTTAAAACAAAGGTAGCTT h36b_ai2.seq

     TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT Majority
              9610      9620      9630      9640      9650
9552 TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT 2603_ai2.seq
9552 TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT 18rs21_ai2.seq
9460 TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT 515_ai2.seq
9449 TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT cjb111_ai2.seq
9487 TTGAAAGTGGTTTGTTCTGGTCATCTGTCTTTTTAACAACTAACTTTCCT h36b_ai2.seq

     TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG Majority
              9660      9670      9680      9690      9700
9602 TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG 2603_ai2.seq
9602 TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG 18rs21_ai2.seq
9510 TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG 515_ai2.seq
9499 TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG cjb111_ai2.seq
9537 TTAGCACCATTTTCCGGTACGGTACTTTCCCCTAAAACATTGGTATTAAG h36b_ai2.seq

     CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT Majority
              9710      9720      9730      9740      9750
9652 CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT 2603_ai2.seq
9652 CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT 18rs21_ai2.seq
9560 CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT 515_ai2.seq
9549 CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT cjb111_ai2.seq
9587 CGGTATTTGCGACAAACAAAAAAGACTTAACGTCAATATTTTAGAAAATT h36b_ai2.seq

     TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT Majority
              9760      9770      9780      9790      9800
9702 TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT 2603_ai2.seq
9702 TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT 18rs21_ai2.seq
9610 TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT 515_ai2.seq
9599 TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT cjb111_ai2.seq
9637 TTTGGTATTTTCTCATTTTACAACTCCTATTGTGCCGAAATGTCGTTTCT h36b_ai2.seq

     AAATCTAAGATCAGATACAGAATATCCTAGAATATACAAACTATCACTTA Majority
              9810      9820      9830      9840      9850
9752 AAATCTAAGATCAGATACAGAATATCCTAGAATATACAAACTATCACTTA 2603_ai2.seq
9752 AAATCTAAGATCAGATACAGAATATCCTAGAATATACAAACTATCACTTA 18rs21_ai2.seq
9660 AAATCTAAGATCAGATACAGAATATCCTAGAATATAAAAACTATCACTTA 515_ai2.seq
9649 AAATCTAAGATCAGATACAGAATATCCTAGAATATACAAACTATCACTTA cjb111_ai2.seq
9687 AAATCTAAGATCAGATACAGAATATCCTAGAATATACAAACTATCACTTA h36b_ai2.seq

     TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAATGTTTTT Majority
              9860      9870      9880      9890      9900
9802 TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAATGTTTTT 2603_ai2.seq
9802 TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAATGTTTTT 18rs21_ai2.seq
9710 TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAAAGTTTTT 515_ai2.seq
9699 TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAATGTTTTT cjb111_ai2.seq
9737 TTATGATATCAATAATTTCTTATTATAAGGTATGGAATTTTAAAGTTTTT h36b_ai2.seq
```

FIGURE 19U

```
      TCCCAATTTTTGAATAATTTTTCTTTTTATTTGATAATCTTATTTTTTAT Majority
               9910      9920      9930      9940      9950
9852  TCCCAATTTTTGAATAATTTTTCTTTTTATTTGATAATCTTATTTTTTAT 2603_ai2.seq
9852  TCCCAATTTTTGAATAATTTTTCTTTTTATTTGATAATCTTATTTTTTAT 18rs21_ai2.seq
9760  TCCCAATTTTTGAATGATTTTTCTTTTTATTTGATA--CATTATTGTCTC 515_ai2.seq
9749  TCCCAATTTTTGAATAATTTTTCTTTTTATTTGATAATCTTATTTTTTAT cjb111_ai2.seq
9787  TCCCAATTTTTGAATGATTTTTCTTTTTATTTGATAATCTTATTTTTTAT h36b_ai2.seq

      TATCTTAGAAATATTTCAATGAGCTTAAGTAGTTGATTTTTCTTTTTTTA Majority
               9960      9970      9980      9990     10000
9902  TATCTTAGAAATATTTCAATGAGCTTAAGTAGTTGATTTTTCTTTTTTTA 2603_ai2.seq
9902  TATCTTAGAAATATTTCAATGAGCTTAAGTAGTTGATTTTTCTTTTTTTA 18rs21_ai2.seq
9808  TATCATGGAAACGTATTAATTA--TAATGTAGTTGTCCCTTACTTACCCC 515_ai2.seq
9799  TATCTTAGAAATATTTCAATGAGCTTAAGTAGTTGATTTTTCTTTTTTTA cjb111_ai2.seq
9837  TATCTTAGAAATACTTCAATTAGCTTAAGTAGTTGATTTTTCTTTTTTTA h36b_ai2.seq

      TGTTTTAAAATATTGCTTAAAAATAATGTTTGAGAGAG--TTTACTGAAT Majority
              10010     10020     10030     10040     10050
9952  TGTTTTAAAATATTGCTTAAAAATAATGTTTGAGAGAG--TTTACTGAAT 2603_ai2.seq
9952  TGTTTTAAAATATTGCTTAAAAATAATGTTTGAGAGAG--TTTACTGAAT 18rs21_ai2.seq
9856  T-TTTCATTGAACTACAAAATATATTTATCTATGATAC--ATTACAATAT 515_ai2.seq
9849  TGTTTTAAAATATTGCTTAAAAATAATGTTTGAGAGAG--TTTACTGAAT cjb111_ai2.seq
9887  TGTTTTAAAATATTGCTTAAAAATAATGTTTGAGAGAGAGTTTACTGAAT h36b_ai2.seq

      TGATTGAAAATTATTTAGAAAAAGACATCCTTAATCAAATAAAACTTCTA Majority
              10060     10070     10080     10090     10100
10000 TGATTGAAAATTATTTAGAAAAAGACATCCTTAATCAAATAAAACTTCTA 2603_ai2.seq
10000 TGATTGAAAATTATTTAGAAAAAGACATCCTTAATCAAATAAAACTTCTA 18rs21_ai2.seq
9903  TGTGTGAAAAATGTTATTATAGATTCCTTTAAAATTAATTTAAAAAAGAC 515_ai2.seq
9897  TGATTGAAAATTATTTAGAAAAAGACATCCTTAATCAAATAAAACTTCTA cjb111_ai2.seq
9937  TGATTGAAAATTATTTAGAAAAAGACATCCTTAATCAAATAAAACTTCTA h36b_ai2.seq

      ACTTTATGCTATGATTACTACCCTTCCATTACTCTAGACAAATCATGTCA Majority
              10110     10120     10130     10140     10150
10050 ACTTTATGCTATGATTACTACCCTTCCATTACTCTAGACAAATCATGTCA 2603_ai2.seq
10050 ACTTTATGCTATGATTACTACCCTTCCATTACTCTAGACAAATCATGTCA 18rs21_ai2.seq
9953  GATATAGACTATATTTTCTTTGCTTACCTTACTTCCCATAACTCTTTTTC 515_ai2.seq
9947  ACTTTATGCTATGATTACTACCCTTCCATTACTCTAGACAAATCATGTCA cjb111_ai2.seq
9987  ACTTTATGCTATGATTACTACCCTTCCATTACTCTAGACAAATCATGTCA h36b_ai2.seq

      TCAACTTGGTTTATCTGAACTACTTATTAGGAA-ATATTGTCATGATTTA Majority
              10160     10170     10180     10190     10200
10100 TCAACTTGGTTTATCTGAACTACTTATTAGGAA-ATATTGTCATGATTTA 2603_ai2.seq
10100 TCAACTTGGTTTATCTGAACTACTTATTAGGAA-ATATTGTCATGATTTA 18rs21_ai2.seq
10003 TAATCCAAATTGGACTGAGAAGCGTATCGATAATGTAATAGCTATTTTCG 515_ai2.seq
9997  TCAACTTGGTTTATCTGAACTACTTATTAGGAA-ATATTGTCATGATTTA cjb111_ai2.seq
10037 TCAACTTGGTTTATCTGAACTACTTATTAGGAA-ATATTGTCATGATTTA h36b_ai2.seq

      ACAACTTTATTTAACAGTCAACTCTCTCTGAATATCGAAAAGAGTACAAT Majority
              10210     10220     10230     10240     10250
10149 ACAACTTTATTTAACAGTCAACTCTCTCTGAATATCGAAAAGAGTACAAT 2603_ai2.seq
10149 ACAACTTTATTTAACAGTCAACTCTCTCTGAATATCGAAAAGAGTACAAT 18rs21_ai2.seq
10053 AAAATTATCCCAAATTCCAAAAATTATTATAGTCACTCAAAGATGCCCTT 515_ai2.seq
10046 ACAACTTTATTTAACAGTCAACTCTCTCTGAATATCGAAAAGAGTACAAT cjb111_ai2.seq
10086 ACAACTTTATTTAACAGTCAACTCTCTCTGAATATCGAAAAGAGTACAAT h36b_ai2.seq

      AGTCTACCAATCTAATGGTGTAACTAGAGAACAAGCTTTCAAATATATTT Majority
              10260     10270     10280     10290     10300
10199 AGTCTACCAATCTAATGGTGTAACTAGAGAACAAGCTTTCAAATATATTT 2603_ai2.seq
10199 AGTCTACCAATCTAATGGTGTAACTAGAGAACAAGCTTTCAAATATATTT 18rs21_ai2.seq
10103 CCCTTATCTGAGTCCTATCATGACGAGTTGGTAAAAGTTGCTAT-CATTT 515_ai2.seq
10096 AGTCTACCAATCTAATGGTGTAACTAGAGAACAAGCTTTCAAATATATTT cjb111_ai2.seq
10136 AGTCTACCAATCTAATGGTGTAACTAGAGAACAAGCTTTCAAATATATTT h36b_ai2.seq

      ATCATCAATCACACGTTTTACAACTTTTAAAATTTTTGATCACGAATGAT Majority
              10310     10320     10330     10340     10350
10249 ATCATCAATCACACGTTTTACAACTTTTAAAATTTTTGATCACGAATGAT 2603_ai2.seq
10249 ATCATCAATCACACGTTTTACAACTTTTAAAATTTTTGATCACGAATGAT 18rs21_ai2.seq
10152 TTTCCCGAACATTTATTTTAGGA---TTAAATCAATTAATCTC---TGAA 515_ai2.seq
10146 ATCATCAATCACACGTTTTACAACTTTTAAAATTTTTGATCACGAATGAT cjb111_ai2.seq
10186 ATCATCAATCACACGTTTTACAACTTTTAAAATTTTTGATCACGAATGAT h36b_ai2.seq
```

FIGURE 19V

```
TCAGGCAGGTTACCTTTAACTTACTTTAGTGAAAAATTTGGACTATCTTG Majority
          10360     10370     10380     10390     10400
10299 TCAGGCAGGTTACCTTTAACTTACTTTAGTGAAAAATTTGGACTATCTTG 2603_a12.seq
10299 TCAGGCAGGTTACCTTTAACTTACTTTAGTGAAAAATTTGGACTATCTTG 18rs21_a12.seq
10196 ACAATTTCATTTCCTTCATGGAACTATCATAGACATGATAAATTAACTAC 515_a12.seq
10196 TCAGGCAGGTTACCTTTAACTTACTTTAGTGAAAAATTTGGACTATCTTG cjb111_a12.seq
10236 TCAGGCAGGTTACCTTTAACTTACTTTAGTGAAAAATTTGGACTATCTTG h36b_a12.seq

TGCAACTGCTTATCGCATACGAAAACATATTAGTCCATTACTAGAAAAAC Majority
          10410     10420     10430     10440     10450
10349 TGCAACTGCTTATCGCATACGAAAACATATTAGTCCATTACTAGAAAAAC 2603_a12.seq
10349 TGCAACTGCTTATCGCATATGAAAACATATTAGTCCATTACTAGAAAAAC 18rs21_a12.seq
10246 TAT-----TCTCCGACCGATAATTACAAATTGGTTAAGTGAAATTGGAGA 515_a12.seq
10246 TGCAACTGCTTATCGCATACGAAAACATATTAGTCCATTACTAGAAAAAC cjb111_a12.seq
10286 TGCAACTGCTTATCGCATACGAAAACATATTAGTCCGTTACTAGAAAAAC h36b_a12.seq

TTGGATTTCAGATTGTCAAAAATACTATTACCGGTG-ACGAGTATCGAAT Majority
          10460     10470     10480     10490     10500
10399 TTGGATTTCAGATTGTCAAAAATACTATTACCGGTG-ACGAGTATCGAAT 2603_a12.seq
10399 TTGGATTTCAGATTGTCAAAAATACTATTACCGGTG-ACGAGTATCGAAT 18rs21_a12.seq
10291 ATACACGTTTAAGGAACAACATTTCTTCTCCTTTGTGCTCACCTAGAAA 515_a12.seq
10296 TTGGATTTCAGATTGTCAAAAATACTATTACCGGTG-ACGAGTATCGAAT cjb111_a12.seq
10336 TTGGATTTCAGATTGTCAAAAATACTATTACCGGTG-ACGAGTATCGAAT h36b_a12.seq

TCGCTATTTAATCGCATTTTTAAATGCTCAATTTGGTATAGAAGTTTATC Majority
          10510     10520     10530     10540     10550
10448 TCGCTATTTAATCGCATTTTTAAATGCTCAATTTGGTATAGAAGTTTATC 2603_a12.seq
10448 TCGCTATTTAATCGCATTTTTAAATGCTCAATTTGGTATAGAAGTTTATC 18rs21_a12.seq
10341 GAATTATCAAAAATCATATTCCTCCGATACAAATAG--CCGTACTAACTA 515_a12.seq
10345 TCGCTATTTAATCGCATTTTTAAATGCTCAATTTGGTATAGAAGTTTATC cjb111_a12.seq
10385 TCGCTATTTAATCGCACATTTAAATGCTCGATTTGGTATAGAAGTTTATC h36b_a12.seq

CCATGTCTAAGATGGATAAATTGCTCATCAAACGATTGTTATTAGAACAC Majority
          10560     10570     10580     10590     10600
10498 CCATGTCTAAGATGGATAAATTGCTCATCAAACGATTGTTATTAGAACAC 2603_a12.seq
10498 CCATGTCTAAGATGGATAAATTGCTCATCAAACGATTGTTATTAGAACAC 18rs21_a12.seq
10389 CAGACTTTATTAATAACAAAATTTTAACAGAATGTTTATCACAGAGGTTT 515_a12.seq
10395 CCATGTCTAAGATGGATAAATTGCTCATCAAACGATTGTTATTAGAACAC cjb111_a12.seq
10435 CCTTATCTAAGATGGATAAATTGCTTATCAAACGATTGTTATTAGAATAC h36b_a12.seq

TCAACTACTTTTACTGCTTCTCATTACTTCC-CAAATACATTTATTTTCT Majority
          10610     10620     10630     10640     10650
10548 TCAACTACTTTTACTGCTTCTCATTACTTCC-CAAATACATTTATTTTCT 2603_a12.seq
10548 TCAACTACTTTTACTGCTTCTCATTACTTCC-CAAATACATTGATTTTCT 18rs21_a12.seq
10439 TCTTCTAAACAGATTCATTTCCACCCTTACTATTTATTAACTGATGATCT 515_a12.seq
10445 TCAACTACTTTTACTGCTTCTCATTACTTCC-CAAATACATTTATTTTCT cjb111_a12.seq
10485 TCAACTACTTTTACTGCTTCTCATTACTTCC-CMAATACATTTATTTTCT h36b_a12.seq

TTGATACATTGTTGTCTCTATCATGGAAACGTATTAATTATAATGTAGTT Majority
          10660     10670     10680     10690     10700
10597 TTGATACATTGTTGTCTCTATCATGGAAACGTATTAATTATAATGTAGTT 2603_a12.seq
10597 TTGATACATTGTTGTCTCTATCATGGAAACGTATTAATTATAATGTAGTT 18rs21_a12.seq
10489 TTCCAATATTACTAATCTTAATCCAGATAT-TATTATTACCAACCCAAAG 515_a12.seq
10494 TTGATACATTGTTGTCTCTATCATGGAAACGTATTAATTATAATGTAGTT cjb111_a12.seq
10534 TTGATACATTATTGTCTCTATCATGGAAACGTATTAATTATAATGTAGTT h36b_a12.seq

GTCCCTTACT---CATCCCTTTTCACTGAACTACAAAATATATTTATCTA Majority
          10710     10720     10730     10740     10750
10647 GTCCCTTACT---CATCCCTTTTCACTGAACTACAAAATATATTTATCTA 2603_a12.seq
10647 GTCCCTTACT---CATCCCTTTTCACTGAACTACAAAATATATTTATCTA 18rs21_a12.seq
10538 CTTTCTTCCTTTATAAAACATGAGATTTCTTCAGAGAGTTTAATTACATA 515_a12.seq
10544 GTCCCTTACT---CATCCCTTTTCACTGAACTACAAAATATATTTATCTA cjb111_a12.seq
10584 GTCCCTTACT---CACCCCTTTTCATTGAACTACAAAATATATTTATCTA h36b_a12.seq

TGATACATTACAATATTGTGTCAAAAATGTTATTATAGATTCCTTTAAAA Majority
          10760     10770     10780     10790     10800
10694 TGATACATTACAATATTGTGTCAAAAATGTTATTATAGATTCCTTTAAAA 2603_a12.seq
10694 TGATACATTACAATATTGTGTCAAAAATGTTATTATAGATTCCTTTAAAA 18rs21_a12.seq
10588 TATTGATTTGATTCATACTCCAGACCAGATCAATGAAA-T-C-C-A-A-G-A-AAT-TA 515_a12.seq
10591 TGATACATTACAATATTGTGTCAAAAATGTTATTATAGATTCCTTTAAAA cjb111_a12.seq
10631 TGATACATTACAATATTGTGTCAAAAATGTTATTATAGATTCCTTTAAAA h36b_a12.seq
```

FIGURE 19W

```
      TTAATTTAAAAAAGACGATATAGACTATATTTTTCTTGC---------T Majority
               10810     10820     10830     10840     10850
10744 TTAATTTAAAAAAGACGATATAGACTATATTTTTCTTGC---------T 2603_a12.seq
10744 TTAATTTAAAAAAGACGATATAGACTATATTTTTCTTGC---------T 18rs21_a12.seq
10638 TTTCATGAATACAAGAAGAAAAAATATTGTAAACTTTTTGCAAAGAACAAT 515_a12.seq
10641 TTAATTTAAAAAAGACGATATAGACTATATTTTTCTTGC---------T cjb111_a12.seq
10681 TTAATTTAAAAAAGACGATATAGACTATATTTTTCTTGC---------T h36b_a12.seq

      TACCTTACTTCCCATAACTCTTTTTCTAATCCAAATTGGAC--TGAGAAG Majority
               10860     10870     10880     10890     10900
10785 TACCTTACTTCCCATAACTCTTTTTCTAATCCAAATTGGAC--TGAGAAG 2603_a12.seq
10785 TACCTTACTTCCCATAACTCTTTTTCTAATCCAAATTGGAC--TGAGAAG 18rs21_a12.seq
10688 GAAATAACTACCCGTAGCTCGTATAACGCTTAAAAATTAACATTAAAAG 515_a12.seq
10682 TACCTTACTTCCCATAACTCTTTTTCTAATCCAAATTGGAC--TGAGAAG cjb111_a12.seq
10722 TACCTTACTTCCCATAACTCTTTTTCTAATCCAAATTGGAC--TGAGAAG h36b_a12.seq

      CGTATCGATAATGTAATA-GCTATTTTCGAAAATTATCCCAAATTCCAAA Majority
               10910     10920     10930     10940     10950
10833 CGTATCGATAATGTAATA-GCTATTTTCGAAAATTATCCCAAATTCCAAA 2603_a12.seq
10833 CGTATCGATAATGTAATA-GCTATTTTCGAAAATTATCCCAAATTCCAAA 18rs21_a12.seq
10738 CTAGAGCATTACACAATGCTCTAGCTTTTTAATGTAATTTTTTTGAATA 515_a12.seq
10730 CGTATCGATAATGTAATA-GCTATTTTCGAAAATTATCCCAAATTCCAAA cjb111_a12.seq
10770 CGTATCGATAATGTAATA-GCTATTTTCGAAAATTATCCCAAATTCCAAA h36b_a12.seq

      AATTATTACAGCCACTCAAAGATGCTCTTCC--CTTATCTGGCTCCTATC Majority
               10960     10970     10980     10990     11000
10882 AATTATTACAGCCACTCAAAGATGCTCTTCC--CTTATCTGGCTCCTATC 2603_a12.seq
10882 AATTATTACAGCCACTCAAAGATGCTCTTCC--CTTATCTGGCTCCTATC 18rs21_a12.seq
10788 AATATAATCCAACTTTTCAACTGTTTTTTTCCAGGTGAAATGTTCTTTAAT 515_a12.seq
10779 AATTATTACAGCCACTCAAAGATGCTCTTCC--CTTATCTGGCTCCTATC cjb111_a12.seq
10819 AATTATTACAGCCACTCAAAGATGCTCTTCC--CTTATCTGGCTCCTATC h36b_a12.seq

      ATGATGAGTTGGTAAAAGTTGCTATCTTTTTTTTCCGAACATTTATTTTA Majority
               11010     11020     11030     11040     11050
10930 ATGATGAGTTGGTAAAAGTTGCTATCTTTTTTT-CCGAACATTTATTTTA 2603_a12.seq
10930 ATGATGAGTTGGTAAAAGTTGCTATCTTTTTTTTCCGAACATTTATTTTA 18rs21_a12.seq
10838 TCTTTTAGCAAAATTCTGTTGTAGTTTCTCTGTTAAATGCCTTATCTTTTA 515_a12.seq
10827 ATGATGAGTTGGTAAAAGTTGCTATCTTTTTTTTCCGAACATTTATTTTA cjb111_a12.seq
10867 ATGATGAGTTGGTAAAAGTTGCTATCTTTTTTTTCCGAACATTTATTTTA h36b_a12.seq

      GGATTAAATCAATTAATCCCTGAAACAATT--TCATTTCCTTCATGGAAC Majority
               11060     11070     11080     11090     11100
10979 GGATTAAATCAATTAATCCCTGAAACAATT--TCATTTCCTTCATGGAAC 2603_a12.seq
10980 GGATTAAATCAATTAATCCCTGAAACAATT--TCATTTCCTTCATGGAAC 18rs21_a12.seq
10888 CTAATAAATCAAGAGATTCATGGAGTGAGTGAGTATTTTCTTCCATGATG 515_a12.seq
10877 GGATTAAATCAATTAATCCCTGAAACAATT--TCATTTCCTTCATGGAAC cjb111_a12.seq
10917 GGATTAAATCAATTAATCCCTGAAACAATT--TCATTTCCTTCATGGAAC h36b_a12.seq

      TATCATAGACATGATAAATTAACTACTATTCTCCGACCGATAATTACAAA Majority
               11110     11120     11130     11140     11150
11027 TATCATAGACATGATAAATTAACTACTATTCTCCGACCGATAATTACAAA 2603_a12.seq
11028 TATCATAGACATGATAAATTAACTACTATTCTCCGACCGATAATTACAAA 18rs21_a12.seq
10938 AATCCTAACTCAGGCTATCGAAAACT-TCAACTGTTCCAGCGGCGATCTG 515_a12.seq
10925 TATCATAGACATGATAAATTAACTACTATTCTCCGACCGATAATTACAAA cjb111_a12.seq
10965 TATCATAGACATGATAAATTAACTACTATTCTCCGACCGATAATTACAAA h36b_a12.seq

      TTGGTTAAGTAAAATTGGAGAATA---CACGTTTAAAGAACAACATTTTC Majority
               11160     11170     11180     11190     11200
11077 TTGGTTAAGTAAAATTGGAGAATA---CACGTTTAAAGAACAACATTTTC 2603_a12.seq
11078 TTGGTTAAGTAAAATTGGAGAATG---CACGTTTAAAGAACAACATTTTC 18rs21_a12.seq
10987 TTGCAATAATAGCACTTGAAAGTAGACCAGCTTCTAAAAATAGAGGTTTGGT 515_a12.seq
10975 TTGGTTAAGTAAAATTGGAGAATA---CACGTTTAAAGAACAACATTTTC cjb111_a12.seq
11015 GTGGTTAAGTGAAATTGGAGAATA---CACATTTAAAGAACAACATTTTC h36b_a12.seq

      TTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATATTCCTCCG Majority
               11210     11220     11230     11240     11250
11124 TTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATATTCCTCCG 2603_a12.seq
11125 TTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATATTCCTCCG 18rs21_a12.seq
11037 AAATCCTGTGGAAACATTGAAGGGAAAACAAAGATATCAGTCTGTGCCA 515_a12.seq
11022 TTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATATTCCTCCG cjb111_a12.seq
11062 TTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATATTCCTCCG h36b_a12.seq
```

FIGURE 19X

```
      A T A C A G A - - T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A T T T T A Majority
                        11260                11270               11280                11290               11300
11174 A T A C A G A - - T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A T T T T A 2603_a12.seq
11175 A T A C A G A - - T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A T T T T A 18rs21_a12.seq
11086 T T A A A G A C A T A G T C T G T T C A A A G T T T A A T T T T C C C A A A A A G T T A A T C T G T 515_a12.seq
11072 A T A C A G A - - T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A T T T T A cjb111_a12.seq
11112 A T A C A G A - - T A G C C G T A C T A A C T A C A G A C T T T A T T A A T A A C C A A A T T T T A h36b_a12.seq

      A C A G A A T G C T T A T T A C A G A G A T T T T C T T C T A A A A A G A T T C A T T T C C A C C C Majority
                        11310                11320               11330                11340               11350
11222 A C A G A A T G C T T A T T A C A G A G A T T T T C T T C T A A A A A G A T T C A T T T C C A C C C 2603_a12.seq
11223 A C A G A A T G C T T A T T A C A G A G A T T T T C T T C T A A A A A G A T T C A T T T C C A C C C 18rs21_a12.seq
11136 T T G G A C T G A T A T T T C T - - - - - C T T T C A A A T G T G C T A A T T C A G G T C C G T C T 515_a12.seq
11120 A C A G A A T G C T T A T T A C A G A G A T T T T C T T C T A A A A A G A T T C A T T T C C A C C C cjb111_a12.seq
11160 A C A G A A T G C T T A T T A C A G A G A T T T T C T T C T A A A A A G A T T C A T T T C C A C C C h36b_a12.seq

      T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A T C C A G Majority
                        11360                11370               11380                11390               11400
11272 T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A T C C A G 2603_a12.seq
11273 T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A T C C A G 18rs21_a12.seq
11181 C C T G C A A T C T G T A A A T A A A C A T T T T C A G A G T A C T G - T G A C A T C A A - - - A A 515_a12.seq
11170 T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A T C C A G cjb111_a12.seq
11210 T T A C T A T T T A T T A A C T G A T G A T C T T T C C A A T A T T A C T A A T C T T A A T C C A G h36b_a12.seq

      A C A T T A T T A T T A C C A A T A A A A A G C T T T C T C C C T T T A T C A A A C A T G A G A T T Majority
                        11410                11420               11430                11440               11450
11322 A C A T T A T T A T T A C C A A T A A A A A G C T T T C T C C C T T T A T C A A A C A T G A G A T T 2603_a12.seq
11323 A C A T T A T T A T T A C C A A T A A A A A G C T T T C T C C C T T T A T C A A A C A T G A G A T T 18rs21_a12.seq
11227 A T G C T T C T A A G A G C A A T T C A A T G C T T T T T C T T A A T - - - - - - - - A A T T C 515_a12.seq
11220 A C A T T A T T A T T A C C A A T A A A A A G C T T T C T C C C T T T A T C A A A C A T G A G A T T cjb111_a12.seq
11260 A C A T T A T T A T T A C C A A T A A A A A G C T T T C T C C C T T T A T C A A A C A T G A G A T T h36b_a12.seq

      T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T C T G A T T A A T A C T T C A G A C C A Majority
                        11460                11470               11480                11490               11500
11372 T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T C T G A T T A A T A C T T C A G A C C A 2603_a12.seq
11373 T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T C T G A T T A A T A C T T C A G A C C A 18rs21_a12.seq
11269 T A C C A G C A T A A G T G A T G A A A A T A T C A T T A G C A G A T T T T T C A A G G T A A G C C 515_a12.seq
11270 T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T C T G A T T A A T A C T T C A G A C C A cjb111_a12.seq
11310 T C T T C A G A G A G T T T A A T T A C A T A T A T T G A T C T G A T T A A T A C T T C A G A C C A h36b_a12.seq

      A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G G A A G A A A A A T A T T Majority
                        11510                11520               11530                11540               11550
11422 A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G G A A G A A A A A T A T T 2603_a12.seq
11423 A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G G A A G A A A A A T A T T 18rs21_a12.seq
11319 G T G T C A G C A A A A T C A G A G C C T A G A C T T T C A G A T A C C G A A T T A T A A A T A A C 515_a12.seq
11320 A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G G A A G A A A A A T A T T cjb111_a12.seq
11360 A A T C A A T C A A A T C C A A A A A A T T A T T T C A T C A A T A C A G G A A G A A A A A T A T T h36b_a12.seq

      G C A A A C T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T C C T - - - - - Majority
                        11560                11570               11580                11590               11600
11472 G C A A A C T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T C C T - - - - - 2603_a12.seq
11473 G C A A A C T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T C C T - - - - - 18rs21_a12.seq
11369 T C C T T T A G C T T C T A T A T T A A A A T G T T T T A A C C A T T C A A A C G C T T C T C T T G G 515_a12.seq
11370 G C A A A C T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T G C T - - - - - cjb111_a12.seq
11410 G C A A A C T T T T T G C A A A A A C T A A T G A A A T A A C T A A T C G T A G C T C C T - - - - - h36b_a12.seq

      A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T - A G A G C A T T G T G T A A T G C T C Majority
                        11610                11620               11630                11640               11650
11517 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T - A G A G C A T T G T G T A A T G C T C 2603_a12.seq
11518 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T - A G A G C A T T G T G T A A T G C T C 18rs21_a12.seq
11419 A T A C C G C A T A A A A A T C T G G A C G A T A A T G C T T A A C A C G C G C T G T G A G A - - - 515_a12.seq
11415 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T - A G A G C A T T G T G T A A T G C T C cjb111_a12.seq
11455 A T A A C T C T T A A A A A T T A A C A T T A A A A A G C T - A G A G C A T T G T G T A A T G C T C h36b_a12.seq

      T A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T Majority
                        11660                11670               11680                11690               11700
11566 T A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T 2603_a12.seq
11567 T A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T 18rs21_a12.seq
11466 - A G A T G T T T A T A G A T A G C T C C A A A G A A A T C T A A A A A A C G A - T T A T T G A C A 515_a12.seq
11464 T A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T cjb111_a12.seq
11504 T A G C T T T T T A A T G T T A A T T T T T T T T G A A T A A T A T A A T C C A A C T T T T C A A C T h36b_a12.seq
```

FIGURE 19Y

```
      GTTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTG Majority
               11710     11720     11730     11740     11750
11616 GTTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTG 2603_al2.seq
11617 GTTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTG 18rs21_al2.seq
11514 GAAAAATGACTTGACCCATGGTCTAAAACAATACTAGGTAAATGGTGCTT 515_al2.seq
11514 GTTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTG cjb111_al2.seq
11554 GTTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTG h36b_al2.seq

      TAGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCAT Majority
               11760     11770     11780     11790     11800
11666 TAGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCAT 2603_al2.seq
11667 TAGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCAT 18rs21_al2.seq
11564 CTTTGCAAAAGATAGCCCTTCTAACGTTGTTAACTGAAAACGTGTATTAC 515_al2.seq
11564 TAGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCAT cjb111_al2.seq
11604 TAGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCAT h36b_al2.seq

      GGAGT---GACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTA Majority
               11810     11820     11830     11840     11850
11716 GGAGT---GACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTA 2603_al2.seq
11717 GGAGT---GACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTA 18rs21_al2.seq
11614 AAATCACAAAATCAATATTTTCATCTGAAACATATTTCATCAGCGTGTTG 515_al2.seq
11614 GGAGT---GACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTA cjb111_al2.seq
11654 GGAGT---GACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTA h36b_al2.seq

      T------CAATAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGC Majority
               11860     11870     11880     11890     11900
11763 T------CAATAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGC 2603_al2.seq
11764 T------CAATAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGC 18rs21_al2.seq
11664 TATTCTCGATTTTTGTTAATAATAGGATAGCGCTGCTTGACAATATTTTT 515_al2.seq
11661 T------CAATAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGC cjb111_al2.seq
11701 T------CAATAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGC h36b_al2.seq

      ACTTGAAAGTAGACCAGCTTCTAAA-----ATAGAGGTTGGTAATCCCTC Majority
               11910     11920     11930     11940     11950
11807 ACTTGAAAGTAGACCAGCTTCTAAA-----ATAGAGGTTGGTAATCCCTC 2603_al2.seq
11808 ACTTGAAAGTAGACCAGCTTCTAAA-----ATAGAGGTTGGTAATCCCTC 18rs21_al2.seq
11714 GGTCGGTAAACGGTAAATTTTTCTACCCTTGTCTTCATCTATAATCGGT. 515_al2.seq
11705 ACTTGAAAGTAGACCAGCTTCTAAA-----ATAGAGGTTGGTAATCCCTC cjb111_al2.seq
11745 ACTTGAAAGTAGACCAGCTTCTAAA-----ATAGAGGTTGGTAATCCCTC h36b_al2.seq

      TGGATACATTGAAGGGTAAACAAAGATATCAGTCTGT-GCCATTAAAGAC Majority
               11960     11970     11980     11990     12000
11852 TGGATACATTGAAGGGTAAACAAAGATATCAGTCTGT-GCCATTAAAGAC 2603_al2.seq
11853 TGGATACATTGAAGGGTAAACAAAGATATCAGTCTGT-GCCATTAAAGAC 18rs21_al2.seq
11763 AAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAAC 515_al2.seq
11750 TGGATACATTGAAGGGTAAACAAAGATATCAGTCTGT-GCCATTAAAGAC cjb111_al2.seq
11790 TGGATACATTGAAGGGTAAACAAAGATATCAGTCTGT-GCCATTAAAGAC h36b_al2.seq

      ATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAATCTGTTTGGACTG Majority
               12010     12020     12030     12040     12050
11901 ATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAATCTGTTTGGACTG 2603_al2.seq
11902 ATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAATCTGTTTGGACTG 18rs21_al2.seq
11813 TAAATCTGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGG 515_al2.seq
11799 ATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAATCTGTTTGGACTG cjb111_al2.seq
11839 ATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAATCTGTTTGGACTG h36b_al2.seq

      ATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTA Majority
               12060     12070     12080     12090     12100
11951 ATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTA 2603_al2.seq
11952 ATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTA 18rs21_al2.seq
11863 GTAGATAATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATA 515_al2.seq
11849 ATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTA cjb111_al2.seq
11889 ATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTA h36b_al2.seq

      AATAAACATTTTCA-GAGTACTGTGACATCGAAAATGCTTCTAAGAGCAA Majority
               12110     12120     12130     12140     12150
12001 AATAAACATTTTCA-GAGTACTGTGACATCGAAAATGCTTCTAAGAGCAA 2603_al2.seq
12002 AATAAACATTTTCA-GAGTACTGTGACATCGAAAATGCTTCTAAGAGCAA 18rs21_al2.seq
11913 TTTATTCACTTTCATTAATAAGCCATCTTTTAAGCCTTTAATCATAGCAA 515_al2.seq
11899 AATAAACATTTTCA-GAGTACTGTGACATCGAAAATGCTTCTAAGAGCAA cjb111_al2.seq
11939 AATAAACATTTTCA-GAGTACTGTGACATCGAAAATGCTTCTAAGAGCAA h36b_al2.seq
```

FIGURE 19Z

```
      TTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGTGATGAAAATAT Majority
               12160     12170     12180     12190     12200
12050 TTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGTGATGAAAATAT 2603_ai2.seq
12051 TTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGTGATGAAAATAT 18rs21_ai2.seq
11963 CTCAATTTTTTTGCTCTTTGGCTCTTCTGCTACAACACTCGAACAAATAT 515_ai2.seq
11948 TTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGTGATGAAAATTC cjb111_ai2.seq
11988 TTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGTGATGAAAATAT h36b_ai2.seq

      CATCAGCAGATTTTTCAAGGTAAGCCGTACCAGCAAAATCAGAGCCTAGA Majority
               12210     12220     12230     12240     12250
12100 CATCAGCAGATTTTTCAAGGTAAGCCGTACCAGCAAAATCAGAGCCTAGA 2603_ai2.seq
12101 CATCAGCAGATTTTTCAAGGTAAGCCGTACCAGCAAAATCAGAGCCTAGA 18rs21_ai2.seq
12012 AATTCGCAGAAATACAAAA-AATTGTATC-GCTTCTTCTTAGCATATT 515_ai2.seq
11998 CATCAGCAGATTTTTCAAGGTAAGCCGTACCAGCAAAATCAGAGCCTAGA cjb111_ai2.seq
12038 CATCAGCAGATTTTTCAAGGTAAGCCGTACTAGCAAAATCAGAGCCTAGA h36b_ai2.seq

      CTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCTATATTAAAATG Majority
               12260     12270     12280     12290     12300
12150 CTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCTATATTAAAATG 2603_ai2.seq
12151 CTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCTATATTAAAATG 18rs21_ai2.seq
12059 TTTTTAAAATATAGATTGCA-TTGCGGATTCAGGTAATATTTTCGAAATT 515_ai2.seq
12048 CTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCTATATTAAAATG cjb111_ai2.seq
12088 CTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCTATATTAAAATG h36b_ai2.seq

      TTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAAATCTGGACGAT Majority
               12310     12320     12330     12340     12350
12200 TTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAAATCTGGACGAT 2603_ai2.seq
12201 TTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAAATCTGGACGAT 18rs21_ai2.seq
12106 GGTGATTGATTCAATAC-ATGAAAAACAGGCCAAAATTTTTAACTCGT 515_ai2.seq
12098 TTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAAATCTGGACGAT cjb111_ai2.seq
12138 TTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAAATCTGGACGAT h36b_ai2.seq

      AATGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAGCTCCAAAGAAA Majority
               12360     12370     12380     12390     12400
12250 AATGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAGCTCCAAAGAAA 2603_ai2.seq
12251 AATGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAGCTCCAAAGAAA 18rs21_ai2.seq
12154 GAAGAGTGTCCAATTTCGTG-TAACAGACCAATAAATTAACCTGAAAG 515_ai2.seq
12148 AATGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAGCTCCAAAGAAA cjb111_ai2.seq
12188 AATGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAGCTCCAAAGAAA h36b_ai2.seq

      TCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCATGGTCTAAAAC Majority
               12410     12420     12430     12440     12450
12300 TCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCATGGTCTAAAAC 2603_ai2.seq
12301 TCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCATGGTCTAAAAC 18rs21_ai2.seq
12203 TCTTATATTCCATCTCTGACAGA-CGATAAATCATTTCAGAGTCAACAAA 515_ai2.seq
12198 TCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCATGGTCTAAAAC cjb111_ai2.seq
12238 TCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCATGGTCTAAAAC h36b_ai2.seq

      AATACTAGGTAAATGGTGATTCTTTGCAAA---AGATAGCCCTTCTAGCG Majority
               12460     12470     12480     12490     12500
12350 AATACTAGGTAAATGGTGATTCTTTGCAAA---AGATAGCCCTTCTAGCG 2603_ai2.seq
12351 AATACTAGGTAAATGGTGATTCTTTGCAAA---AGATAGCCCTTCTAGCG 18rs21_ai2.seq
12252 ATCAATAAAACA------TCTCTTCTGCAAAGCC-AGATGTTTCCTTCGAAA 515_ai2.seq
12248 AATACTAGGTAAATGGTGATTCTTTGCAAA---AGATAGCCCTTCTAGCG cjb111_ai2.seq
12288 AATACTAGGTAAATGGTGCTTCTTTGCAAA---AGATAGCCCTTCTAGCG h36b_ai2.seq

      TTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATATTTTCATCT Majority
               12510     12520     12530     12540     12550
12397 TTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATATTTTCATCT 2603_ai2.seq
12398 TTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATATTTTCATCT 18rs21_ai2.seq
12296 CGCTCGTTTTCATTAAAGCAGTCGAAGTAATACACTCTTCAATTTCTTTA 515_ai2.seq
12295 TTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATATTTTCATCT cjb111_ai2.seq
12335 TTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATATTTTCATCT h36b_ai2.seq

      GAAACATATTTCATCAGCGTGTTGTATTCTCGATT-----TTTGTTAATA Majority
               12560     12570     12580     12590     12600
12447 GAAACATATTTCATCAGCGTGTTGTATTCTCGATT-----TTTGTTAATA 2603_ai2.seq
12448 GAAACATATTTCATCAGCGTGTTGTATTCTCGATT-----TTTGTTAATA 18rs21_ai2.seq
12346 TAGTCAAATTCTTGCATCACTAAATTTTCACGGTTCATATCTTGATACAA 515_ai2.seq
12345 GAAACATATTTCATCAGCGTGTTGTATTCTCGATT-----TTTGTTAATA cjb111_ai2.seq
12385 GAAACATATTTCATCAGCGTGTTGTATTCTCGATT-----TTTGTTAATA h36b_ai2.seq
```

FIGURE 19AA

```
      ATAGGATAGC------------GCTGCTTGACAATATTTTTGGTCGGTAA Majority
                12610     12620     12630     12640     12650
12492 ATAGGATAGC------------GCTGCTTGACAATATTTTTGGTCGGTAA 2603_al2.seq
12493 ATAGGATAGC------------GCTGCTTGACAATATTTTTGGTCGGTAA 18rs21_al2.seq
12396 ATAAGATAACATACTGATCTTAGTTAAATGAAGGTAATTTTCATAATTAT 515_al2.seq
12390 ATAGGATAGC------------GCTGCTTGACAATATTTTTGGTCGGTAA cjb111_al2.seq
12430 ATAGGATAGC------------GCTGCTTGACAATATTTTTGGTCGGTAA h36b_al2.seq

      ACGGTAAATTTTTC----TACCCTTGTCTTCATCTATAATCGGTAAATCA Majority
                12660     12670     12680     12690     12700
12530 ACGGTAAATTTTTC----TACCCTTGTCTTCATCTATAATCGGTAAATCA 2603_al2.seq
12531 ACGGTAAATTTTTC----TACCCTTGTCTTCATCTATAATCGGTAAATCA 18rs21_al2.seq
12446 CTATCAAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAA 515_al2.seq
12428 ACGGTAAATTTTTC----TACCCTTGTCTTCATCTATAATCGGTAAATCA cjb111_al2.seq
12468 ACGGTAAATTTTTC----TACCCTTGTCTTCATCTATAATCGGTAAATCA h36b_al2.seq

      CCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAACCAAATC Majority
                12710     12720     12730     12740     12750
12576 CCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAACCAAATC 2603_al2.seq
12577 CCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAACCAAATC 18rs21_al2.seq
12496 TCAAATTCTTGTGTTACTGCAAATTGACCGATACAGTTCAAAGCATATGC 515_al2.seq
12474 CCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAACCAAATC cjb111_al2.seq
12514 CCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTTAACCAAATC h36b_al2.seq

      TGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGAT Majority
                12760     12770     12780     12790     12800
12626 TGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGAT 2603_al2.seq
12627 TGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGAT 18rs21_al2.seq
12546 AATCCCTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCAT 515_al2.seq
12524 TGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGAT cjb111_al2.seq
12564 TGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGAT h36b_al2.seq

      AATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATATTTATC Majority
                12810     12820     12830     12840     12850
12676 AATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATATTTATC 2603_al2.seq
12677 AATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATATTTATC 18rs21_al2.seq
12596 TATAATCGGCTACTAATTGAGAAATTTCTTCCTTATTTTTCG-----AGC 515_al2.seq
12574 AATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATATTTATC cjb111_al2.seq
12614 AATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCCATATTTATC h36b_al2.seq

      CACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAGCAACTATTT Majority
                12860     12870     12880     12890     12900
12726 CACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAGCAACTATTT 2603_al2.seq
12727 CACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAGCAACTATTT 18rs21_al2.seq
12641 CATTATCTACGATA--------TAGATGTGGCTTACTTGAGGATAAATTG 515_al2.seq
12624 CACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAGCAACTATTT cjb111_al2.seq
12664 CACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAGCAACTATTT h36b_al2.seq

      TTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAATTCATTTCGC Majority
                12910     12920     12930     12940     12950
12776 TTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAATTCATTTCGC 2603_al2.seq
12777 TTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAATTCATTTCGC 18rs21_al2.seq
12683 TTTGAATGTTCTGATCT-AAGCGTTCAATATTGGGGTTAA--AGGTGACA 515_al2.seq
12674 TTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAATTCATTTCGC cjb111_al2.seq
12714 TTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAATTCATTTCGC h36b_al2.seq

      ATAAATACTAAATATTT-GTATCGCTTCTTCTTACCATATTTTTTTATAA Majority
                12960     12970     12980     12990     13000
12826 ATAAATACTAAATATTT-GTATCGCTTCTTCTTACCATATTTTTTTATAA 2603_al2.seq
12827 ATAAATACTAAATATTT-GTATCGCTTCTTCTTACCATATTTTTTTATAA 18rs21_al2.seq
12730 ATACCCGCTAAATATTTTATGTTCTATGCTCTTTTCTAAAATCTCTAAAT 515_al2.seq
12724 ATAAATACTAAATATTT-GTATCGCTTCTTCTTACCATATTTTTTTATAA cjb111_al2.seq
12764 ATAAATACTAAATATTT-GTATCGCTTCTTCTTACCATATTTTTTTATAA h36b_al2.seq

      TATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTGAATGATTC Majority
                13010     13020     13030     13040     13050
12875 TATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTGAATGATTC 2603_al2.seq
12876 TATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTGAATGATTC 18rs21_al2.seq
12780 AACTGAATGACTGGTGCTTTG-GTTATAAAAACGATACCGACATAGAT-- 515_al2.seq
12773 TATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTGAATGATTC cjb111_al2.seq
12813 TATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTGAATGATTC h36b_al2.seq
```

FIGURE 19AB

```
      AATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAAT  Majority
          13060     13070     13080     13090     13100
12925 AATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAAT  2603_a12.seq
12926 AATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAAT  18rs21_a12.seq
12827 AGTTACTGCTACTAAACTTTGAATGACATAATTTACCAATGATACTGACA  515_a12.seq
12823 AATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAAT  cjb111_a12.seq
12863 AATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAAT  h36b_a12.seq

      TTCGTGTAACAGACCAATAAAATTAA--CCTGATAAGTCTTATATCCCAT  Majority
          13110     13120     13130     13140     13150
12975 TTCGTGTAACAGACCAATAAAATTAA--CCTGATAAGTCTTATATCCCAT  2603_a12.seq
12976 TTCGTGTAACAGACCAATAAAATTAA--CCTGATAAGTCTTATATCCCAT  18rs21_a12.seq
12877 TTTGAGTATTGATATAATAGAGTACAGCTCCACTAAGAGTAGCAGCAATT  515_a12.seq
12873 TTCGTGTAACAGACCAATAAAATTAA--CCTGATAAGTCTTATATCCCAT  cjb111_a12.seq
12913 TTCGTGTAAAAGACCAATAAAATTAA--CCTGATAAGTCTTATATCCCAT  h36b_a12.seq

      CTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAAACATCT  Majority
          13160     13170     13180     13190     13200
13023 CTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAAACATCT  2603_a12.seq
13024 CTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAAACATCT  18rs21_a12.seq
12927 AAATAGCGCAGTATTCCTCTTGTTA-ATTCTTTAAAA---GTAAATACAT  515_a12.seq
12921 CTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAAACATCT  cjb111_a12.seq
12961 CTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAAACATCT  h36b_a12.seq

      CTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATTAAAGCA  Majority
          13210     13220     13230     13240     13250
13073 CTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATTAAAGCA  2603_a12.seq
13074 CTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATTAAAGCA  18rs21_a12.seq
12973 CTTTTAAAGAGATAGTTTGATATAGGGAGAACAATAAATTCAGTAATAACT  515_a12.seq
12971 CTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATTAAAGCA  cjb111_a12.seq
13011 CTTCTGCAAAGCCAGATGTTTCTTCAAAAACGCTCGTTTTCATTAAAGCA  h36b_a12.seq

      GCCGAAGTAATACACTCTTCAATTTCT-----TTATAGTCAAATTCT-TG  Majority
          13260     13270     13280     13290     13300
13123 GCCGAAGTAATACACTCTTCAATTTCT-----TTATAGTCAAATTCT-TG  2603_a12.seq
13124 GCCGAAGTAATACACTCTTCAATTTCT-----TTATAGTCAAATTCT-TG  18rs21_a12.seq
13023 GTAGAGATAATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATT  515_a12.seq
13021 GCCGAAGTAATACACTCTTCAATTTCT-----TTATAGTCAAATTCT-TG  cjb111_a12.seq
13061 GCCGAAGTAATACACTCTTCAATTTCT-----TTATAGTCAAATTCT-TG  h36b_a12.seq

      CATCACTAAATTT-TCACGGTTCATATCTTGATACAAACAAGATAACATA  Majority
          13310     13320     13330     13340     13350
13167 CATCACTAAATTT-TCACGGTTCATATCTTGATACAAACAAGATAACATA  2603_a12.seq
13168 CATCACTAAATTT-TCACGGTTCATATCTTGATACAAACAAGATAACATA  18rs21_a12.seq
13073 AAGCACAACATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTT  515_a12.seq
13065 CATCACTAAATTT-TCACGGTTCATATCTTGATACAAACAAGATAACATA  cjb111_a12.seq
13105 CATAACTAAATCT-TCACGGTTCATATCTTGATACAAACAAGATAACATA  h36b_a12.seq

      C----CGACCTTAGGTAAATGAAGGTAATTTTCATAA----TTATCTATC  Majority
          13360     13370     13380     13390     13400
13216 C----CGACCTTAGGTAAATGAAGGTAATTTTCATAA----TTATCTATC  2603_a12.seq
13217 C----CGACCTTAGGTAAATGAAGGTAATTTTCATAA----TTATCTATC  18rs21_a12.seq
13123 TTGTACGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAA  515_a12.seq
13114 C----CGACCTTAGGTAAATGAAGGTAATTTTCATAA----TTATCTATC  cjb111_a12.seq
13154 C----CGACCTTAGGTAAAAGAAGGTAATTTTCATAA----TTATCTATC  h36b_a12.seq

      AAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAA  Majority
          13410     13420     13430     13440     13450
13258 AAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAA  2603_a12.seq
13259 AAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAA  18rs21_a12.seq
13173 GAAATGCAAATGATCATCAAAGACTCAACAGCCATTAGCGGACCGACAGC  515_a12.seq
13156 AAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAA  cjb111_a12.seq
13196 AAATCACCTAGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAA  h36b_a12.seq

      TTCTTGTGCTACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATTC  Majority
          13460     13470     13480     13490     13500
13308 TTCTTGTGCTACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATTC  2603_a12.seq
13309 TTCTTGTGCTACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATCC  18rs21_a12.seq
13223 TTCATACTCTTTCCCAAA--AAAGAAGACAG-CAAATGTCGACGAAACTC  515_a12.seq
13206 TTCTTGTGCTACTGCAAATTGACCAATACAGTTCAAAGCATATGCAATTC  cjb111_a12.seq
13246 TTCTTGTGCTACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATTC  h36b_a12.seq
```

FIGURE 19AC

```
      CTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAA Majority
               13510     13520     13530     13540     13550
13358 CTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAA 2603_ai2.seq
13359 CTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAA 18rs21_ai2.seq
13270 CTACCACTTGCCCC--AATCATTAACATTGAGATAGCATTGTTAAATCAA 515_ai2.seq
13256 CTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAA cjb111_ai2.seq
13296 CTTTATTTTCTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAA h36b_ai2.seq

      TCGGCTACTAATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTAC Majority
               13560     13570     13580     13590     13600
13408 TCGGCTACTAATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTAC 2603_ai2.seq
13409 TCGGCTACTAATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTAC 18rs21_ai2.seq
13318 CAAGTTCAAAGTAAGCCTTAGCTTGCTTTCTTTTTCTGAAGAAAAT     515_ai2.seq
13306 TCGGCAACTAATTGAGAAATTTCCTCCTTATTTTTCGAGCCATTATCTAC cjb111_ai2.seq
13346 TCGGCAACTAATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTAC h36b_ai2.seq

      GATGTAGATATGGCTTACTTGAGGA-----TAAATTGCTCGAATGTTCTG Majority
               13610     13620     13630     13640     13650
13458 GATATAGATGTGGCTTACTTGAGGA-----TAAATTGCTCGAATGTTCTG 2603_ai2.seq
13459 GATATAGATGTGGCTTACTTGAGGA-----TAAATTGCTCGAATGTTCTG 18rs21_ai2.seq
13368 AGACTGCAGAGCCTTGGTAAGAAAACTCCTCCGATTGCTGAAACAATGGT 515_ai2.seq
13356 GATGTAGATATGGCTTACTTGAGGA-----TAAATTGCTCGAATGTTCTG cjb111_ai2.seq
13396 GATGTAGATATGGCTTACTTGAGGA-----TAAATTGCTCGAATGTTCTG h36b_ai2.seq

      ATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAATACCCGCTAAATATT Majority
               13660     13670     13680     13690     13700
13503 ATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAATACCCGCTAAATATT 2603_ai2.seq
13504 ATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAATACCCGCTAAATATT 18rs21_ai2.seq
13418 AATAAAATACGAATAATTTTATCAGATTGATCAAAATATCCAGCACTGA  515_ai2.seq
13401 ATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAATACCCGCTAAATATT cjb111_ai2.seq
13441 ATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAATACCCGCTAAATATT h36b_ai2.seq

      TCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAACTGAATGAC----- Majority
               13710     13720     13730     13740     13750
13553 TCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAACTGAATGAC----- 2603_ai2.seq
13554 TCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAACTGAATGAC----- 18rs21_ai2.seq
13468 GCACACTATGCATTGCTCCCAGCATACTCTTGTTAAGAGAGAGATAAATC 515_ai2.seq
13451 TCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAACTGAATGAC----- cjb111_ai2.seq
13491 TCATGTTCTGTGCTCTTTTCTAAAATCTCTAAATAACTGAATGAC----- h36b_ai2.seq

      ----TGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTACTGCT Majority
               13760     13770     13780     13790     13800
13598 ----TGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTACTGCT 2603_ai2.seq
13599 ----TGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTACTGCT 18rs21_ai2.seq
13518 TGCATGGTGATTTGG-GGTAAAAATAACATTAAAGTTGGTCTTAGATGTA 515_ai2.seq
13496 ----TGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTACTGCT cjb111_ai2.seq
13536 ----TGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTACTGCT h36b_ai2.seq

      ACTAAACTTTGAAT----GACATAATTTACCAATGATACTGACATTTGAG Majority
               13810     13820     13830     13840     13850
13644 ACTAAACTTTGAAT----GACATAATTTACCAATGATACTGACATTTGAG 2603_ai2.seq
13645 ACTAAACTTTGAAT----GACATAATTTACCAATGATACTGACATTTGAG 18rs21_ai2.seq
13567 TTAAGATATCTAATCTACTAAATGGTATCTTGATGATTTCATGCTTTAAA 515_ai2.seq
13542 ACTAAACTTTGAAT----GACATAATTTACCAATGATACTGACATTTGAG cjb111_ai2.seq
13582 ACTAAACTTTGAAT----GACATAATTTACCAATGATACTGACATTTGAG h36b_ai2.seq

      TATTGATATAATAGAGTACAGCTCCACTAAGA-GTAGCAGCAATTAAATA Majority
               13860     13870     13880     13890     13900
13690 TATTGATATAATAGAGTACAGCTCCACTAAGA-GTAGCAGCAATTAAATA 2603_ai2.seq
13691 TATTGATATAATAGAGTACAGCTCCACTAAGA-GTAGCAGCAATTAAATA 18rs21_ai2.seq
13617 TGATGCTAGACTGTTAAATTTCCGAGTAATGATGCAAATGCAATTAAA-- 515_ai2.seq
13588 TATTGATATAATAGAGTACAGCTCCACTAAGA-GTAGCAGCAATTAAATA cjb111_ai2.seq
13628 TATTGATATAATAAAGTACAGCTCCACTAAGA-GTAGCAGCAATTAAATA h36b_ai2.seq

      GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG Majority
               13910     13920     13930     13940     13950
13739 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG 2603_ai2.seq
13740 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG 18rs21_ai2.seq
13665 ----AACATATAAAGTGCTAGGTCATCTGGTGTCTTCACAACTAAGAAGG 515_ai2.seq
13637 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG cjb111_ai2.seq
13677 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG h36b_ai2.seq
```

FIGURE 19AD

```
      AGATAGCTTGATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA Majority
               13960     13970     13980     13990     14000
13789 AGATAGCTTGATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA
13790 AGATAGCTTGATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA 2603_al2.seq
13711 TCAAAC..........AACTGTAATAAGCTTAACAATGGTATTACCAATG 18rs21_al2.seq
13687 AGATAGCTTGATATAGAGAGACAATAAATTCAGTAATAACTGTAGAGATA 515_al2.seq
13727 AGATAGCTTGATAGAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA cjb111_al2.seq
                                                         h36b_al2.seq

      ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC Majority
               14010     14020     14030     14040     14050
13839 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC
13840 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC 2603_al2.seq
13752 ACAATTATCTTAAAGTCTTCAACTTCAATAAAATACCACGAAATATCAGT 18rs21_al2.seq
13737 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC 515_al2.seq
13777 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC cjb111_al2.seq
                                                         h36b_al2.seq

      ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTA----AGCTTTTGTA Majority
               14060     14070     14080     14090     14100
13889 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTA----AGCTTTTGTA
13890 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTA----AGCTTTTGTA 2603_al2.seq
13802 TGCAGTCGCTAATAAATTAATACCTTGAATTAAATAGAATAATTGCCATT 18rs21_al2.seq
13787 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTA----AGCTTTTGTA 515_al2.seq
13827 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTA----AGCTTTTGTA cjb111_al2.seq
                                                         h36b_al2.seq

      CGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAAT Majority
               14110     14120     14130     14140     14150
13935 CGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAAT
13936 CGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAAT 2603_al2.seq
13852 CGTTTGTAAGTAAGACAAATCCAAAGAATAATAACGTTGCTATAGATAACA 18rs21_al2.seq
13833 CGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAAT 515_al2.seq
13873 CGTCTTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAAT cjb111_al2.seq
                                                         h36b_al2.seq

      GCAAATGATCATCAAAGACTCAAC-------------ACCGACAACTTCAT Majority
               14160     14170     14180     14190     14200
13985 GCAAATGATCAT............-------------..............
13986 GCAAATGATCAT............-------------.............. 2603_al2.seq
13902 GAGGCAAGTTTTAATACAACTAAT-------------TCCCAAAATATCTT 18rs21_al2.seq
13883 GCAAATGATCATCAAAGACTCAACAGCCATTAGCGGACCGACAACTTCAT 515_al2.seq
13923 GCAAATGATCATCAAAGACTCAACAGCCATTAGCGGACCGACAGCTTCAT cjb111_al2.seq
                                                         h36b_al2.seq

      ACCCTTTCCCAAAAAAGAAGAC-GCA-----AAXXXXXXXXXXXXXXXXXX Majority
               14210     14220     14230     14240     14250
13997 ......................-.CA------AA
13998 ......................-.CA------A                   2603_al2.seq
13940 ACTTTCTTTTATACTTATTTGACTGCA-----CAAAAGCTATCTCACGATT 18rs21_al2.seq
13933 ACCCTTTCCCAAAAAAGAAGACAGCAAATGTTGATGAAACTCCTACCACT  515_al2.seq
13973 ACCCTTTCCCAAAAAAGAAGACGGCA------AA                  cjb111_al2.seq
                                                         h36b_al2.seq

      XXXXXXXXXXXXXXXXXX                                 Majority
               14260
14000
14000                                                    2603_al2.seq
13985 ACCGTA--CAAAGATATC                                 18rs21_al2.seq
13983 CCCGCAATCATTAACATT                                 515_al2.seq
14000                                                    cjb111_al2.seq
                                                         h36b_al2.seq
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 19AE

```
    GTTGCTTCATAAGTTGTCTGAAGGCCTAATTCTAAAGTCACATGCATTCT Majority
             10        20        30        40        50
1   G................................................. 2603_ai2.seq
1   GTTGCTTCATAAGTTGTCTGAAGGCCTAATTCTAAAGTCACATGCATTCT nem316_ai2.seq

    TTCAGAAAGTTCAGCGAGATAGTATATAGTTTCATCAGGTAAGCAATCCG Majority
             60        70        80        90       100
2   .................................................. 2603_ai2.seq
51  TTCAGAAAGTTCAGCGAGATAGTATATAGTTTCATCAGGTAAGCAATCCG nem316_ai2.seq

    GCCTTGTTCCGATGTTGATTCCGATAACTCCTGGCTCATTAATAGCCTGT Majority
            110       120       130       140       150
2   GCCTTGTTCCGATGTTGATTCCGATAACTCCTGGCTCATTAATAGCCTGT 2603_ai2.seq
101 GCCTTGTTCCGATGTTGATCCCGATAACTCCTGGCTCATTAATAGCCTGT nem316_ai2.seq

    TCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAATT Majority
            160       170       180       190       200
52  TCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAATT 2603_ai2.seq
151 TCGTAACGCTCTTTAATTATCTCTAACTTAGCATGGGTATTGGTAAAATT nem316_ai2.seq

    TTGAAAATAGACTAAGTATTTATTAACCTCGGGCCACTTTCTATGCATGA Majority
            210       220       230       240       250
102 TTGAAAATAGACTAAGTATTTATTAACCTCAGGCCACTTTCTATGCATGA 2603_ai2.seq
201 TTGAAAATAGACTAAGTATTTATTAACCTCGGGCCACTTTCTATGCATGA nem316_ai2.seq

    AATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAACT Majority
            260       270       280       290       300
152 AATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAACT 2603_ai2.seq
251 AATCAATTTCTTTATAGAATTGTTCACGAATAGGAGCTTCTGGAGCAACT nem316_ai2.seq

    ATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCATCCTCCTCTAGC Majority
            310       320       330       340       350
202 ATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCACCCTCCTCTAGC 2603_ai2.seq
301 ATAGCATCCCCTGAACCAGAAACTGTGCAAAAAGTGCATCCTCCTCTAGC nem316_ai2.seq

    AACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAATT Majority
            360       370       380       390       400
252 AACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAATT 2603_ai2.seq
351 AACTGTTCCGTCTCTGTTAGGACAGTCAAAACCAGCATCTATAGGTAATT nem316_ai2.seq

    TAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACGA Majority
            410       420       430       440       450
302 TAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACGA 2603_ai2.seq
401 TAAATATTTTTTCTCCAAAGAGTTCTCGATAATAATCATTAATCGCACGA nem316_ai2.seq

    TAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACCT Majority
            460       470       480       490       500
352 TAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACCT 2603_ai2.seq
451 TAACGTTTTTTCATAGGATAATTGTATCACAATTTTAACTAAAATAACCT nem316_ai2.seq

    CACTACTACAATAAAACTTAAAAAGATTGGAACGTCAGTTAGTCCCAATC Majority
            510       520       530       540       550
402 CACTACTACAATAAAACTAAAAAAGATTGGAACGTCAGTTAGTCCCAATC 2603_ai2.seq
501 CACTACTACAATAAAACTTAAAAAGATTGGAACGTCAGTTAGTCCCAATC nem316_ai2.seq

    TTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATACG Majority
            560       570       580       590       600
452 TTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATACG 2603_ai2.seq
551 TTTTATTTACTTCACTTTCTTTAACCAATCCTTGGCTAAAAAGATATACG nem316_ai2.seq

    CAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATCT Majority
            610       620       630       640       650
502 CAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATCT 2603_ai2.seq
601 CAGTTAGATTCAAAATACCATAAGCAAGTATAAAACCAGCTAAAACATCT nem316_ai2.seq
```

Figure 20

GTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAAT Majority

660 670 680 690 700

552 GTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAAT 2603_ai2.seq

651 GTCGGAAAATGAACCCCTAGGTAAATACGAGATAACCCAATTAAAAAAAT nem316_ai2.seq

GAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGCA Majority

710 720 730 740 750

602 GAGCAAACCCAAAGTACCTTGGCACAACAGTTTCCATATACTCTTAGGCA 2603_ai2.seq

701 GAGCAAACCCAAAATACCTTGGCACAACAGTTTCCATATACTCTTAGGCA nem316_ai2.seq

TATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCCC Majority

760 770 780 790 800

652 TATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCCC 2603_ai2.seq

751 TATAGTACTGCAATAAAATAATAATACTCCCAAATATCATAAATGTTCCC nem316_ai2.seq

ATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT Majority

810 820 830 840 850

702 ATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT 2603_ai2.seq

801 ATCGAGTGCCCACTGGGAAACGAATAGCCACCTGCAAATACTAAATGGGT nem316_ai2.seq

TAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACATA Majority

860 870 880 890 900

752 TAAAGTTGGTCTTACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACATA 2603_ai2.seq

851 TAAAGTTGGTCTCACTCTTTGAAAAATAAGTTTTAAAGAAAGTATACATA nem316_ai2.seq

TACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCAC Majority

910 920 930 940 950

802 TACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCAC 2603_ai2.seq

901 TACCAGAGATAATAGCATTTACTGCGATAAATCTAGCTTGAGGATACCAC nem316_ai2.seq

TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT Majority

960 970 980 990 1000

852 TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT 2603_ai2.seq

951 TTCTTAAGGTAACAGAAAGTGACGCTCATAATCGCAATAGCTATCTGGCT nem316_ai2.seq

TACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGAT Majority

1010 1020 1030 1040 1050

902 TACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGAT 2603_ai2.seq

1001 TACAGTATTACCAATCACAGTGATTAACTTGAAAAATCTTGTAGAAAGAT nem316_ai2.seq

TTGGCAACTGTCCTCTAACACTTTCTTGAATGTTTTGGTCAAATGCAATT Majority

1060 1070 1080 1090 1100

952 TTGGCAACTGTCCTCTAACACTTTCTTGAATAGTTTGGTCAAATGAAATT 2603_ai2.seq

1051 TTGGCAACTGTCCTCTAACACTTTCTTGAATGTTTTGGTCAAATGCAATT nem316_ai2.seq

ACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT Majority

1110 1120 1130 1140 1150

1002 ACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT 2603_ai2.seq

1101 ACAGTGTCGGGCCAATATTTGATGACCAATCCTAAACTGAAAAATAAGAT nem316_ai2.seq

AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA Majority

1160 1170 1180 1190 1200

1052 AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA 2603_ai2.seq

1151 AATAGCAATAAATGCTTGAATAAGTTTACTATTTTGACGAGATAACATTA nem316_ai2.seq

GTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA Majority

1210 1220 1230 1240 1250

1102 GTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA 2603_ai2.seq

1201 GTCTTTTTATATCTTTCTAATATTGGCAAACAAGCCACGTAAGTTAGATA nem316_ai2.seq

GAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACCA Majority

1260 1270 1280 1290 1300

1152 GAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACCA 2603_ai2.seq

1251 GAAAACAATCGAAATTAAAATTCCCTCAACGATATTAAATGGAATAACCA nem316_ai2.seq

FIGURE 20A

```
     TTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTTA Majority
              1310      1320      1330      1340      1350
1202 TTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTTA 2603_ai2.seq
1301 TTGTTAAAAGGTAATTGCCTACACCAATAAATGTTCTGATATCAAAGTTA nem316_ai2.seq

     GCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCAT Majority
              1360      1370      1380      1390      1400
1252 GCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCAT 2603_ai2.seq
1351 GCAAATATAGCATACAAAGGAATCGCAAAGACATAGTTGAGAGCTACCAT nem316_ai2.seq

     AGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCTT Majority
              1410      1420      1430      1440      1450
1302 AGATACGGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCTT 2603_ai2.seq
1401 AGATACAGTCAAGCTAACTGTACCAAATAGACTAGCTTTAATAAAATCTT nem316_ai2.seq

     TTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAAA Majority
              1460      1470      1480      1490      1500
1352 TTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAAA 2603_ai2.seq
1451 TTGCACTCTCTCTATTTTTCCAGAAAATAGCGAAACTTGCTAAAAATAAA nem316_ai2.seq

     GCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCACG Majority
              1510      1520      1530      1540      1550
1402 GCTAGAGCAACCATATTCATCGGTAAACCAATAAACGTTTCTGGACCACG 2603_ai2.seq
1501 GCTAGAGCAACCATATTCATCGGTAAACCGATAAAGGTTTCTGGACCACG nem316_ai2.seq

     ATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAAC Majority
              1560      1570      1580      1590      1600
1452 ATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAAC 2603_ai2.seq
1551 ATTAGCAAGTATAACTTTTAAAAGTGATCTTAATAAGAGTACACCATAAC nem316_ai2.seq

     TTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGAA Majority
              1610      1620      1630      1640      1650
1502 TTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGAA 2603_ai2.seq
1601 TTGATTTCAAATCAAATAAAATAAAAGCAACTAACATCGGAAGGATTGAA nem316_ai2.seq

     AAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT Majority
              1660      1670      1680      1690      1700
1552 AAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT 2603_ai2.seq
1651 AAATCAACCTTTAAAAATTCTGCTCCTGGTATTAATGGAAATGAAACCAT nem316_ai2.seq

     CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTAC Majority
              1710      1720      1730      1740      1750
1602 CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTAC 2603_ai2.seq
1701 CATCAATACAAAAGATAAGGCAGAAAGAATGGCGATTGTCACCATTTTAC nem316_ai2.seq

     GTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAGC Majority
              1760      1770      1780      1790      1800
1652 GTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAGC 2603_ai2.seq
1751 GTGTATTTGTCATAAAAAAATTCCTCCAATTTAAATAAATTGAAAGAAGC nem316_ai2.seq

     TCCAAAGGTAAGCGTATGTACGCGAAAAAAACCTTTGTCTTCTCCCATCC Majority
              1810      1820      1830      1840      1850
1702 TCCAAAGGTAAGCGTATGTACGCGAAAAAA-CCTTTGTCTTCTCCCATCC 2603_ai2.seq
1801 TCCAAAGGTAAGCGTAGGTACGCGAAAAAAACCTTTGTCTTCTCCCATCC nem316_ai2.seq

     AGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCGG Majority
              1860      1870      1880      1890      1900
1751 AGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCGG 2603_ai2.seq
1851 AGACTTTACTGTCGGTTGTGGAATCTCACCACATCAGCTTTCGCTCGCGG nem316_ai2.seq

     ACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTCG Majority
              1910      1920      1930      1940      1950
1801 ACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTCG 2603_ai2.seq
1901 ACTGATGCTTCACAACTGACAAATAAGTTGGAAGCGATTACCGCCGGTCG nem316_ai2.seq
```

FIGURE 20B

```
     GGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTTG Majority
               1960      1970      1980      1990      2000
1851 GGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTTG 2603_a12.seq
1951 GGAATTACACCCTGCCCTGAAGACACCTATAGCATAACAAAAAAAACTTG nem316_a12.seq

     CAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA Majority
               2010      2020      2030      2040      2050
1901 CAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA 2603_a12.seq
2001 CAATTGCAAGTTTTTTAATTACTAATTAGTAGTAGTGATTAAAAATCATA nem316_a12.seq

     TTAATACCAAATTACTATGCTGTATCGTTTCTTTCAGATTTGCTATTTTT Majority
               2060      2070      2080      2090      2100
1951 TTAATACCAAATTACTATACTGTATCGTTTCTTTCAGATTTGCTATTTTT 2603_a12.seq
2051 TTAATACCAAATTACTATGCTGTATCGTTTCTTTCAGATTTGCTATTTTT nem316_a12.seq

     AGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGAA Majority
               2110      2120      2130      2140      2150
2001 AGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGAA 2603_a12.seq
2101 AGTTTTTCTTAAAAAGATAAACAAAATTCCCAAAATAATACAACCAAGAA nem316_a12.seq

     TTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT Majority
               2160      2170      2180      2190      2200
2051 TTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT 2603_a12.seq
2151 TTGTCAGTCCTCCACCAATAATCATTCCTGTTTTAGGAAGAAATGATTGT nem316_a12.seq

     GGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTTT Majority
               2210      2220      2230      2240      2250
2101 GGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTTT 2603_a12.seq
2201 GGAAAAAGCGGTTGTGATGGTTTAGGATTTGTTGGTGGAGGAGTTTCTTT nem316_a12.seq

     TTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACTA Majority
               2260      2270      2280      2290      2300
2151 TTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACTA 2603_a12.seq
2251 TTCGTTTTCTACCTCTACTTCCTGTGTTTTATTAGCAACTACAGCAACTA nem316_a12.seq

     CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA Majority
               2310      2320      2330      2340      2350
2201 CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA 2603_a12.seq
2301 CAGCATCCTTCATAGATATACGGTAACCAGTTAGTGCTTTTGCTTCTCGA nem316_a12.seq

     AAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC Majority
               2360      2370      2380      2390      2400
2251 AAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC 2603_a12.seq
2351 AAAATATACTTACCAGGTAATAAACCTTCAACCTCAATTTCTCCCTTATC nem316_a12.seq

     ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCAT Majority
               2410      2420      2430      2440      2450
2301 ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCAT 2603_a12.seq
2401 ATCAGTTACTAATGAAGTAATCCCATCTTGATCGGTCGTAAATCGTCCAT nem316_a12.seq

     TTTTAAAGCGAACTGGCTGATTCTGGTTATCGTATAATACAAATATTACT Majority
               2460      2470      2480      2490      2500
2351 TTTTAAAGCGAACTGGCTGATTCTGGTTATCGTATAATACAAATATTACT 2603_a12.seq
2451 TTTTAAAGCGAACTGGCTGATTCTGGTTATCGTATAATACAAATATTACT nem316_a12.seq

     CCTGATAGCCTTTTCTTTATCTTTCCTTCTTTTGTATATTTAATAAGTTT Majority
               2510      2520      2530      2540      2550
2401 CCGGATAGCCTTTTCTTTATCTTTCCTTCTTTTGTATATTTAATAAGTTT 2603_a12.seq
2501 CCTGATAGCCTTTTCTTTATCTTTCCTTCTTTTGTATATTTAATAAGTTT nem316_a12.seq

     TAATCGGCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTTG Majority
               2560      2570      2580      2590      2600
2451 TAATCGGCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTTG 2603_a12.seq
2551 TAATCGGCCTGTTTCAACTTTTCGCTTAGGATTTATCTGTAATTGATTTG nem316_a12.seq
```

FIGURE 20C

```
     ATAACTTATCATCTGGTATTTCAATATAAAAAGGTACTATTGTTGAAACG Majority
               2610      2620      2630      2640      2650
2501 ATAACTTATCATCTGGTAATTCAATATAAAAAGGTACTATTGTTGAAACG 2603_ai2.seq
2601 ATAACTTATCATCTGGTATTTCAATATAAAAAGGTACTATTGTTGAAACG nem316_ai2.seq

     CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC Majority
               2660      2670      2680      2690      2700
2551 CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC 2603_ai2.seq
2651 CTTTGATCAGCTTTATAAGCACGACCAAAGTACGAACCATTTGGGAGTGC nem316_ai2.seq

     TATCTTTGTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT Majority
               2710      2720      2730      2740      2750
2601 TATCTTTGTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT 2603_ai2.seq
2701 TATCTTTGTCTGACCATTAGTATCAGTAGGAGAAGTCAAGATACTCTTAT nem316_ai2.seq

     ACTTCTGGTTCAATTCGCTATCTGTCATTTGGCTCAATAAATCAACTTTT Majority
               2760      2770      2780      2790      2800
2651 ACTTCTGGTTCAATTCGCTATCTGTCATTTGGCTCAATAAATCAACTTTT 2603_ai2.seq
2751 ACTTCTGGTTCAATTCGCTATCTGTCATTTGGCTCAATAAATCAACTTTT nem316_ai2.seq

     AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT Majority
               2810      2820      2830      2840      2850
2701 AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT 2603_ai2.seq
2801 AAGTTGTCAGTCACAGTCCATAAACGATAAGAAATCCCCTCCTCTGTAGT nem316_ai2.seq

     ATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG Majority
               2860      2870      2880      2890      2900
2751 ATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG 2603_ai2.seq
2851 ATTTGGCTGAAGTCCTATCTGTGTGATTGTTAGTTGATTAGGGGTATCAG nem316_ai2.seq

     CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG Majority
               2910      2920      2930      2940      2950
2801 CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG 2603_ai2.seq
2901 CATTTACACTGGCTACCGAAAAAAACGCTAATTGTACCAATCCTAAAAAG nem316_ai2.seq

     CAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATTT Majority
               2960      2970      2980      2990      3000
2851 CAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATTT 2603_ai2.seq
2951 CAACATAGTAGAAGTCCTAAACTTTTTCTAATCTTTTTCATTTTTGATTT nem316_ai2.seq

     CCCTTTCTTTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAGC Majority
               3010      3020      3030      3040      3050
2901 CCCTTTCTCTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAGC 2603_ai2.seq
3001 CCCTTTCTTTTTCTCTCTTTAAATTTTCGTTTTAAATATAATAGTAAAGC nem316_ai2.seq

     GACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAGT Majority
               3060      3070      3080      3090      3100
2951 GACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAGT 2603_ai2.seq
3051 GACTAATATAAGAATAACTAGGATTGATAAGAGGAAATAAAGTTTATAGT nem316_ai2.seq

     GTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTACA Majority
               3110      3120      3130      3140      3150
3001 GTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTACA 2603_ai2.seq
3101 GTGTTTGCAATTCTTTCATTAAATAGTTCTTTTCTTTAACAGGAGGTACA nem316_ai2.seq

     TACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAATA Majority
               3160      3170      3180      3190      3200
3051 TACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAATA 2603_ai2.seq
3151 TACTTGATTCGATGCCCTCTAACTAGTAAACGATGTGAATTAATCGAATA nem316_ai2.seq

     AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA Majority
               3210      3220      3230      3240      3250
3101 AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA 2603_ai2.seq
3201 AGGTGTACATGTTAGCAAAGTCGCATAATCCTTACCTTTAACAACCAATA nem316_ai2.seq
```

FIGURE 20D

ATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATAG Majority
3260 3270 3280 3290 3300
3151 ATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATAG 2603_a12.seq
3251 ATTTAGAAAAATTATCTGGCTTTACAACACTTATTTGATCAACCTTATAG nem316_a12.seq

GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAAG Majority
3310 3320 3330 3340 3350
3201 GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAAG 2603_a12.seq
3301 GCTAAAACTTCTTTGATATTATGAATATAAAAAATTTTTCCTTTTTTAAG nem316_a12.seq

TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCTG Majority
3360 3370 3380 3390 3400
3251 TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCTG 2603_a12.seq
3351 TTTATCTAAATCTGTAAATAACTTAGCTTTAGGTAAGCCGCGATGAGCTG nem316_a12.seq

TGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTCA Majority
3410 3420 3430 3440 3450
3301 TGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAGGAGGTTCCTTCA 2603_a12.seq
3401 TGATAACAGTATGTGAACTTTTTCCACCAATTGGCAAAGAAGTTCCTTCA nem316_a12.seq

AGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGAT Majority
3460 3470 3480 3490 3500
3351 AGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGAT 2603_a12.seq
3451 AGGTGTCCTGCTCCTTTTTCAAGAACACTACTGGTAGTCCCCGCATAGAT nem316_a12.seq

AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG Majority
3510 3520 3530 3540 3550
3401 AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG 2603_a12.seq
3501 AGGTAATTTTTGCTTGATAGACGGTATATCAATATATCCAATCATTTCAG nem316_a12.seq

TAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTCA Majority
3560 3570 3580 3590 3600
3451 CAATCTCAAGCATGTGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTCA 2603_a12.seq
3551 TAATCTCAAGCATATGGGCGTATTCAGCAATACCTTTTTTTTCTTTTTCA nem316_a12.seq

GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTTT Majority
3610 3620 3630 3640 3650
3501 GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTTT 2603_a12.seq
3601 GTATAGGGATCTGATAGGCGGCTTGGGTCCAGTGTTCTATTATAAGCTTT nem316_a12.seq

TGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTTT Majority
3660 3670 3680 3690 3700
3551 TGCTAACTCAAATCGTCTATTAATCTCTTTAGTATTTAATTTTTGGGTTT 2603_a12.seq
3651 TGCTAACTCAAATCGTCTATTAATCTCTTTAGCATTTAATTTTTGGGTTT nem316_a12.seq

GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA Majority
3710 3720 3730 3740 3750
3601 GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA 2603_a12.seq
3701 GATTATCAAAGTTAGTTACTTGATTATTAGCTTTAATATTATAGTACCAA nem316_a12.seq

TTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATAG Majority
3760 3770 3780 3790 3800
3651 TTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATAG 2603_a12.seq
3751 TTTGAAATAAAAGGATATGAGGTTATCAAAAGACCAACTAAGAACAATAG nem316_a12.seq

TATCAGGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGGT Majority
3810 3820 3830 3840 3850
3701 TATCAAGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGGT 2603_a12.seq
3801 TATCAGGCCTACATTCATCCATCGATTTAAAACGACCGATTTCTTAAGGT nem316_a12.seq

TTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTTG Majority
3860 3870 3880 3890 3900
3751 TTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTTG 2603_a12.seq
3851 TTTTCTGAAATTTTCCTCCCATTATGATTCAATTCCTTTTCTAACACTTG nem316_a12.seq

FIGURE 20E

```
     CTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGCA Majority
             3910      3920      3930      3940      3950
3801 CTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGCA 2603_al2.seq
3901 CTAAACGATTTTTTTGACGTTGACGTTTTATTAACCAAAGTAACCAAGCA nem316_al2.seq

     ATAATAACTAAAGATATATAGAATAGATATCTATAAATCGTGTTTAAATG Majority
             3960      3970      3980      3990      4000
3851 ATAATAACTAAAGATATATAGAATAGATATCTATAAATCGTGTTTAAATG 2603_al2.seq
3951 ATAATAACTAAAGATATATAGAAAAGATATCTATAAATCGTGTTTAAATG nem316_al2.seq

     ACCTTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGAT Majority
             4010      4020      4030      4040      4050
3901 ACCGTCTTTTATTAATTTTTCATCAATAGGACCTTTATAAGGGATACGAT 2603_al2.seq
4001 ACCTTCTTTCATTAATTTTTCATCAATAAGACCTTTATAAGGAATACGAT nem316_al2.seq

     GTCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGTT Majority
             4060      4070      4080      4090      4100
3951 GTCCCCTTACTAAAAGTCTGTGTGTATTGATCATAATCGGGGTGCAAGTC 2603_al2.seq
4051 GTCCCCTTACTAAAAGTCTGTGTGTATTGACCATAATCGGGGTGCAAGTT nem316_al2.seq

     AATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT Majority
             4110      4120      4130      4140      4150
4001 AATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT 2603_al2.seq
4101 AATAAGGTTGCATAATCATGTCCAGGAACAACCAACAAATCTGAAAAGTT nem316_al2.seq

     ATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCTT Majority
             4160      4170      4180      4190      4200
4051 ATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCTT 2603_al2.seq
4151 ATCGGGTGTAACGACTTTTATCTGATCTACTTGATATGCTATCGTTTCTT nem316_al2.seq

     TTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTTA Majority
             4210      4220      4230      4240      4250
4101 TTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTTA 2603_al2.seq
4201 TTATGTTTTGAATATAAAACTTATCTCCTTTTTTTAACTTTTTAAGGTTA nem316_al2.seq

     GAAAAGAGTTCTTTATCTGGAATTCCTGAGTGCGCTGTTATAACGGTATG Majority
             4260      4270      4280      4290      4300
4151 GAAAAGAGTTCTTTATCTGGAATTCCTGAATGCGCTGTTATAACAGTATG 2603_al2.seq
4251 GAAAAGAGTTCTTTATCTGGAATTCCTGAGTGCGCTGTTATAACGGTATG nem316_al2.seq

     TGTGCTATTTCCTCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCTC Majority
             4310      4320      4330      4340      4350
4201 TGTACTATTGCCCCCAATTGGAAGAGAGGTACCTTCTAAATGCCCTGCTC 2603_al2.seq
4301 TGTGCTATTTCCTCCAATTGGAAGAGAAGTACCTTCTAAATGCCCTGCTC nem316_al2.seq

     CTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTGA Majority
             4360      4370      4380      4390      4400
4251 CTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTGA 2603_al2.seq
4351 CTTTAGATAGAACTTCTTGACTTGAACCTGCAAATATAGGGAGTTTTTGA nem316_al2.seq

     CCTATCTTAGGAACTGAAATTGTTCCGATTTTTTCACTTACCTCTAACAT Majority
             4410      4420      4430      4440      4450
4301 CCTATCTTAGGAACTGAAATTGTTCCGATTTTTTCACTTACCTCTAACAT 2603_al2.seq
4401 CCTATCTTAGGAACTGAAATTGTGCCGATTTTTTCACTTACCTCTAACAT nem316_al2.seq

     ACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGAT Majority
             4460      4470      4480      4490      4500
4351 ACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGAT 2603_al2.seq
4451 ACGGGCGTACTCTGCTACCCCCTTTTGAATTCGTTTTTTCTCATAAGGAT nem316_al2.seq

     CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGTC Majority
             4510      4520      4530      4540      4550
4401 CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGCC 2603_al2.seq
4501 CTTCAAGATGGACATTATTTAAAGAATCATTATAAGCTTGTGCTAGAGTC nem316_al2.seq
```

FIGURE 20F

ATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAAA Majority
4560 4570 4580 4590 4600
4451 ATACGTCGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAAA 2603_ai2.seq
4551 ATACGACGATTGATTTCTTTCTGACTAAGTTTTTTAGCAGCTCTCTCAAA nem316_ai2.seq

ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA Majority
4610 4620 4630 4640 4650
4501 ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA 2603_ai2.seq
4601 ATCCTGTGTTTGATTATTAGATTCTATCGTATAGTAAAAACGTGATACCA nem316_ai2.seq

CTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGGA Majority
4660 4670 4680 4690 4700
4551 CTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGGA 2603_ai2.seq
4651 CTGGATACAATAAAATAGATAGACCTATTAGAAAAAGAATGATAAAAGGA nem316_ai2.seq

AGATTTGACTTCTTCTTTTTTTTTTGTTTTTTTGTTGATTTTTTTAGTCTT Majority
4710 4720 4730 4740 4750
4601 AGATTTGACTTCTTCTTTTTTTTTTGTTTTTTTGATGATTTTTTTAGTCTT 2603_ai2.seq
4701 AGATTTGACTTCTTCTTTTTA----GTTTTTTTGTTGATATTTTTAGTCTT nem316_ai2.seq

CACGTCATCTCCTAGATAATGGCTCTTGCTTATGATCTAAGAGTACTTCT Majority
4760 4770 4780 4790 4800
4651 CACGTCATCTCCTAGATAATGGCTCTTGCTTATGATCTAAGAGTACTTCT 2603_ai2.seq
4748 CACGTCATCTCCTAAATAATGGCTCTTGCTTATGATCTAAGAGTACTTCT nem316_ai2.seq

ACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACATC Majority
4810 4820 4830 4840 4850
4701 ACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGCGCTTATACATC 2603_ai2.seq
4798 ACTGAAATACCCTTAGATCATAAGCACAGCTTTAACTGTGCTTATACACC nem316_ai2.seq

ATCAAAGACTAGCCTTAAGCTTCCTTTGATTGGCGTTTTTTTCATGATAAC Majority
4860 4870 4880 4890 4900
4751 ATCAAAGACTAGCCTTAAGCTTCCTTTGATTGGCGTGTTTTTCATGATAAC 2603_ai2.seq
4848 ATCAAAGACTAGCCTTAAGCTTCCTGTGATTGACGTTTTTTTCATGACAAC nem316_ai2.seq

TACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC Majority
4910 4920 4930 4940 4950
4801 TACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC 2603_ai2.seq
4898 TACTGCTCCAAGCATAATGCTTAAACCAATAATTGTGAAAAGAATTGTAC nem316_ai2.seq

CAATACCACCTGTTTGTGGGATTGTTACTTTTTTGTTTTGTACTTGTTTG Majority
4960 4970 4980 4990 5000
4851 CAATACCACCTGTTTGTGGGATTGTTACGTTTTTATTTTGTACACGTGTG 2603_ai2.seq
4948 CAATACCACCTGTTTGTGGGATGGTTACTTTTTTGTTTTGAACTTGTTGG nem316_ai2.seq

GCATCTTTTTTTACAGGTTTTTTGTTATCTGCGTTGTCAGTTTTAGCCCC Majority
5010 5020 5030 5040 5050
4901 GCATCTTTTT------GGTTGGTGTTAGCAACGTAGTCAATGTTACGACC 2603_ai2.seq
4998 GCATCTTTTTTTTACAGATCCTTTATCATATGCGATGTCAGTTGTAGCCCC nem316_ai2.seq

TTTTCTGTATGATGTTTGATTTACTTCAAAGTTTATATTACCTGCCAATT Majority
5060 5070 5080 5090 5100
4945 TGTTATGTATGACCCTTCATTAACTACAAAGTTAATATTACCTGCCAACT 2603_ai2.seq
5048 TTTGCTATATGATGTGGGAGTTACTTCAAAGTTTACATCACCTGACAATG nem316_ai2.seq

TCGCATATCCTGCTGGTGCTTGTGTTTCTTCCAGGTTGTAAGTGCCTTTT Majority
5110 5120 5130 5140 5150
4995 TAGCAAATCCTGCTGGAGCAAGTGTTTCTTCAAGGTTGTAAGTACCGTGT 2603_ai2.seq
5098 TCGCATAACCTGCTGGTGCTTGAGTTTCTTCCAAGGTATAAGTGCCTTTA nem316_ai2.seq

TCCAGACCTGTAATTTCAAATTGACCTTGGTCGTTTGAGGTGTATTTAAT Majority
5160 5170 5180 5190 5200
5045 GCAAGACCTGTAAGTTCAAATTGACCTTGATCGTTTGAAGTGTAGGTAAT 2603_ai2.seq
5148 TCCAAACCAGTAATTTCAAATTGACCACCGGCGTTAGAGATCAATTTAAG nem316_ai2.seq

FIGURE 20G

```
      GGCTTTATCAGCCTTTTTATCTGCTATCCATTCATATGCTTTGTTAGCTT Majority
          5210      5220      5230      5240      5250
5095  GGCTCTAGC------TTATCTGTTATCCAGTCATAAGCTGTACGAGCCT 2603_ai2.seq
5198  AACATTATCAGCCTTTTTATCTGCAAACCATTCATATGAGTAGTTAGCTT nem316_ai2.seq

      TAATGAAGGCTTCATTGTAATCTTTTTGTTTTTTTTTGATAGTTTCTTTT Majority
          5260      5270      5280      5290      5300
5139  CAATGAAGGCTGCATGGTAATCTGGTTGTTTAGTTTTGATAAGTTCTTTT 2603_ai2.seq
5248  TAACAAAAGCGTCATTGTAAGCTTTTTGTTTTTGATTCAACAGTAGCCAAT nem316_ai2.seq

      GCTGTTTTTCCTTTTTGGCCTTTTTGTTTTTTTTTTAGTCAAGTTGTTATA Majority
          5310      5320      5330      5340      5350
5189  GCAGTAATTCCTTTTTCACCTTTTTGGTCTGTTGCAGACAAGTTGTTATA 2603_ai2.seq
5298  GCTGTTTTACCTTCTTGGCCTTCTTGTTTTTCTTTAGTCAAGTCGTTATA nem316_ai2.seq

      AGCTTTGATAGCTTCATCTAATGCTATTTTAGCAGTTTTTGCAGCTGCCT Majority
          5360      5370      5380      5390      5400
5239  AGCAGCGATAGCTTCATCTAAAGCTATTTT------TTAGCAGCTAAAG 2603_ai2.seq
5348  AGCTTTAACAGCTTCATCGAATGCTAGTTTAGCAGTTTTTACAGCTGCCT nem316_ai2.seq

      TTTTTTGAGCTTTTGCTTGATCTGCTTTACGTGCCAGGTATTTTCCTTCT Majority
          5410      5420      5430      5440      5450
5283  TTTTTTGAACTTCTGATTGATCTGCTTTAAGAGCAAGGTATTTACCTGCT 2603_ai2.seq
5398  TTGCTTCAGCAGTTGCTGCAACTGCTTTACGTGCCAAGTATTTTCCTTCT nem316_ai2.seq

      GAGTTTTTCACAAGGAATTTTGCTCCTGCCAAGCGTTCAGTTTTAGCTTG Majority
          5460      5470      5480      5490      5500
5333  GAGTTTTTCACAACGAATTGTGCACCAGCCAAACGGTCACCTTG---TTG 2603_ai2.seq
5448  ---TTGTTAACAAGGAAGGGTAGCTCCTGCCAAGCGTTCAGTGTTAGCTTG nem316_ai2.seq

      ATTTGTTTTGACAAATTTCTTTCCATGAGTTTCCACTTTTGGTTCTGTTG Majority
          5510      5520      5530      5540      5550
5380  ATTAGTTTTGACAAATTTCTTACCATGAGTTTCAACTTTTGGTTCAGTTG 2603_ai2.seq
5495  ATTTGTTTTCACAAATTTACGTCCATAAGTCACCACTTTTGGTTCTGATG nem316_ai2.seq

      GGTTGATTGGTGTTGGGTTATTTGAGTTTTTGTTGTTGTTGATGGTTACT Majority
          5560      5570      5580      5590      5600
5430  GGTTCAATGGTGTTGGGTTATCAGAATCTTTGGTATTGGTAATGGTTACT 2603_ai2.seq
5545  GGTTGATTGGAGTTGGATCATTTGAGTTTTTGTTGTTCTTGATAGTCACA nem316_ai2.seq

      TTACCATTTTTTAGTTTTATTTCTTTTGGAGTGTAGCCGGTAACACGTTC Majority
          5610      5620      5630      5640      5650
5480  TTACCATTTTCTAGATTTATTGCACTTC----GTAACCAGAAACACGTTC 2603_ai2.seq
5595  ACACCATTTTTAAATGATAACATATTCTGGAGTGTAGCCGCTAACACGTTC nem316_ai2.seq

      TGCGATGCTGTATGTTTTGGTATTTTCTAGACCTGTGAATTTACCCGAGT Majority
          5660      5670      5680      5690      5700
5527  TGAGATCATGTATGATTTG---TTTTCTAGACCAGTGAATTTACCCGAGA 2603_ai2.seq
5645  GACAACGCGGTAAGTTTTAGTATTATCCAAACCTGTGAA--------AGT nem316_ai2.seq

      ATTTTCCAGGTGTTTTGGTTTTGTTGCTTTTTGTGAGGTCGCTTTTACCT Majority
          5710      5720      5730      5740      5750
5574  AGTTACCAGATACTTCAAAATTTGATACCATTTCCAAGGTCGATTGTACCT 2603_ai2.seq
5687  ATGTTCAAAGCGTGATGGTTTTGTTGCTTCGTGTGAAGCAACGTTCACCC nem316_ai2.seq

      TTGTATGTTTTTGTTTGTTTTTCTTGCGCGGTAGTTTTATCTTTATCTTT Majority
          5760      5770      5780      5790      5800
5624  TTAGATGTTTTTGTCAATGATACTGAAGCAACAGTTTTATCTTTATCTTT 2603_ai2.seq
5737  ATGTAC--CATCCGTTTGTTTTTCTTGCCAGGTAAAGATAGCTTTAACT--- nem316_ai2.seq

      CAATGTGTAAGCAACGTTTACATCATTAGTTGTACTTTCTTCTGCCCAGT Majority
          5810      5820      5830      5840      5850
5674  CAATGTGTAAACAACGTTTACAGCATCAGGTGCAATTCCGTCAGACCAAG 2603_ai2.seq
5784  ----------GCAACATTAACATCAGTAATTGTACCATCTACTGCCCAGT nem316_ai2.seq
```

FIGURE 20H

TTTTAGTGACTTTTATTTCTTGCTTTCGTGGTGTACCTTGTTGTTGTTTC Majority
5860 5870 5880 5890 5900
5724 TTTTAGCAACTGTTAGTTCACCCTTTGAAGCGTGTAACAAGCAAGTTCAGTC 2603_ai2.seq
5824 CTTTAATGACTTTAATTTCTTGCTTAGCTGGAGTACCTTCTTGTGGTTCA nem316_ai2.seq

CTTTCTTTCGTTGGTTTGTTACCATAGTCCAATTTGATATCATTGGTTTC Majority
5910 5920 5930 5940 5950
5774 AAGTCTTTACCTGGTTTGTTACCATACGACAATTTGATATCATTGGATTC 2603_ai2.seq
5874 CTTTCTTCCGTTGGGTTATTACCATAGTCCAATTTAACATCATTGGTTTC nem316_ai2.seq

TGGATTTTCAATAGTTGTTTGGCCGTTCACCGTAGCTGTGTAAGTGATTT Majority
5960 5970 5980 5990 6000
5824 TGGATTATCAATAATTGCTTGACCATTAACAGTAGCACTATAAGTCAATG 2603_ai2.seq
5924 TGGAACTTCAACAGTAGTGGAGCCGTTCACCGTAGCTGAGTAAGTGATCT nem316_ai2.seq

TGATTTCAATATCTTTTTTTTTTGCTGCTTTTTCTACTTTTGCCAGTCCT Majority
6010 6020 6030 6040 6050
5874 TAAATTCAATATCAGCTGTTTTAGCTGCTTTTTCCAATTTGCCCAATCCA 2603_ai2.seq
5974 TGATTTCAACATCTTTGTGTTTTGCAGCAGCTGCTACTGCTGCAAGACCT nem316_ai2.seq

TTTGCTTTGAATTTTAGTGTGAAACCTTGGTCATCTGTTCTGAGTTTGTA Majority
6060 6070 6080 6090 6100
5924 TCAGCTGTGAATTTTAATGTGAAACCACGGGCATCAATGCTAAGTTCATA 2603_ai2.seq
6024 GTTGCATTCAAGGCAAGACGGAAACCTTGGTCATCTGTTACGAGTTTGTA nem316_ai2.seq

GTTTGTATCCTTAGGAAAATTTTTTGTATTTCCTGTTGCTTTAAGGTTGT Majority
6110 6120 6130 6140 6150
5974 GTCTGTATCCTTAGCAAAAGTTTCTGTAGTTCCTGAAGCTTTAAGGCTAA 2603_ai2.seq
6074 GTTTAAAC----AGGAAAATCTTTACCATCCAATGTTACTTTAAGGTTGT nem316_ai2.seq

TAGTTGAACCCATTGTCAAACCTTTTGTCATTCTATCTGTCCAAACCAGT Majority
6160 6170 6180 6190 6200
6024 CAGTTGAACCCATTGTCAAACCATTTGACATTATATCTGTCCAAACCAAG 2603_ai2.seq
6121 T-GTTGAAC-----GTCAAACCTTTAGTCATGCTATCAGTCCAAACCAGT nem316_ai2.seq

TTTTTGTATTTTGAGCCTTTGTGAATTTTTGTTTTAACTTTGTATTGTTT Majority
6210 6220 6230 6240 6250
6074 TTTTCGTATTTAGAACCTTTGTGAATTTTTGTTTTAACTTCATAAGGAAC 2603_ai2.seq
6165 TTCTTATAGTCTGAGCCTTTAAGAATTTTTGTTGCAACTATGTATTCTTT nem316_ai2.seq

TTTTTTACCGATTTTCGCTGTGGCAGTTGCTTTGTCACGTGCATAATTAC Majority
6260 6270 6280 6290 6300
6124 AACTTTACCGATTTCAGCAGTAGCAGTTGCTTTGTCACGTGCATAATTAC 2603_ai2.seq
6215 TTTGTCACCAACTGTCGCTGAGACAATCACCTTTGTC-------------- nem316_ai2.seq

CATAATTTGCGCCAGCTGTCAAAAGTCTATTTTCGTTTGTCAGTGTTGTC Majority
6310 6320 6330 6340 6350
6174 CATAATTTGCGCCAGCTGTCAAAAGTCTATTAACATCTGTCAATGCTGTC 2603_ai2.seq
6251 ------------------------------TTTTCGGTTGTCAGTATAATC nem316_ai2.seq

AAGTTGTTTGTTTTTGCAAAGTTTTTATCTATTTGTGGTTTTTTTCAGT Majority
6360 6370 6380 6390 6400
6224 AAATCGTTTGTTTAGCAAAGTTTTTATCAATTTCTGGTTTTTCTTCAGT 2603_ai2.seq
6272 AAGATCTTTATC--TGCAAAGTTCTTATCTACTTGTGGTTTTGTTTCAGT nem316_ai2.seq

GTTCTTTGGATAAATGTGGGCATCTTTAACAACACCTTGTTTGTTTACCA Majority
6410 6420 6430 6440 6450
6274 GTTCTTTGGATAAACATGGGCATCAGCAACAACACCATCTTCATTTACCA 2603_ai2.seq
6320 ATTCTTTGGATAAATGTGAGCATCTTTAACAACACCTTCGTTGTTTACCA nem316_ai2.seq

ATGGCAGAGTGATTTTAACTGGAACTGCTTTTGAATCAGCCAGGATGGAA Majority
6460 6470 6480 6490 6500
6324 ATGGAAGAGTGATGTTAACTGGAACGGCTTTTGAAGCAGCCAGGAGGGAA 2603_ai2.seq
6370 ATGGCAGAGTGATTTTAACTGGAACTGCTTTTGAATCAGCCAAGATAGAA nem316_ai2.seq

FIGURE 20I

98/487


```
     CCGTTGTTGTTGTAGTTTGATTTTTCTTTCACTTCAACGATTTTGTACTT Majority
              6510      6520      6530      6540      6550
6374 CCATTATTGTTGTAAGTAGATTTTGATTTAACTTCAACAATTTTAAACTC 2603_a12.seq
6420 CCGTTGTTATGGTAGTTTGATTTTTCTTTCAATTCAACGATTTGGTAAGT nem316_a12.seq

     TCCTTTTAATTTTTTTGGTGTTGAAAGCAAGTCCAGTGTCTTTTTTTGGTG Majority
              6560      6570      6580      6590      6600
6424 GCCTTTCAATCCTTTGGTGTTGAAAACAAGTCCAGTATCTCCCTCTGGTG 2603_a12.seq
6470 TCCTTTTAACTTAGCAGTGTTAAAAGCAAAACCAGTGTCTTTGTTAACC nem316_a12.seq

     TTGATCCAGGCACGGCCTCATCTTTATTTTCTTTTGTTTCCGGAGTATCG Majority
              6610      6620      6630      6640      6650
6474 TCAATCCAGACACGGCCTCATCAATATTTACTGTTATTTCAGGAGTACCA 2603_a12.seq
6520 GTGA---AAGAACAGCACCACCTTCAGCATCTTTAGCTTCCAAAGTATCG nem316_a12.seq

     TCTTTCTTAATTAAGGCTGGTGTTAATTTGTTACCTTCTTTTTCCTTAAT Majority
              6660      6670      6680      6690      6700
6524 TCTTTATTAATTAAGGCTGGTGTTAATTTGTTACCTTCTTTTGCCTTAAC 2603_a12.seq
6567 ACTTCCTTA------------------------------CATTTTCAGTAAT nem316_a12.seq

     GTATTGCATTTTACCAGTTTTATTTTTTTTCAAAGCTAAAGCAAAGAACG Majority
              6710      6720      6730      6740      6750
6574 ATATTGCAGTTTACCAGTTTTATGTTGTTTCAAAGCTAAAGCAAAGAACG 2603_a12.seq
6590 GAATT---TTGTACCAGTTTCATTTTTT----------AAAACAAAGAAAG nem316_a12.seq

     CACCTTTGATTTCTTTAGCTTCGTTTGAGCCAAAGTAAGCTTTAAGGTCA Majority
              6760      6770      6780      6790      6800
6624 CACCTTCGATTTCTTTAGATTCCTC---GCCAAAGTAACCAGCAAGGTCA 2603_a12.seq
6628 CACCCTTAATTTCTTTAGCATCGGTTGAGCCAAAATAAGATTTAAGGTCA nem316_a12.seq

     TTAATTTGTTTACCTTTGTAGTCTTTTTCGTTCTTACCTTTTGTTCCTTG Majority
              6810      6820      6830      6840      6850
6671 GAAATAGGTCCACCTTTGTAGTCTTTTCCGTTAAAGACCTGTAGTTCCTGG 2603_a12.seq
6678 TTAATTTGTTTACCAACATAATCGCTATCATTCTTACCTTTTGTACCTTC nem316_a12.seq

     GGAGTTACTTTTGTTAAGTTTTGCTTGTGTTTTGACAATCTTGTGCAAGG Majority
              6860      6870      6880      6890      6900
6721 GAAGTTACTTTTGTTAAGATTTGATTCGGTTGCAAAATCTTGTGCAAAG 2603_a12.seq
6728 AG---TAAAGTTATTCAAATGCAGCTTGTGGCATGACAATCTTGTGCAAGG nem316_a12.seq

     TCACTGTATTAGTTGTTGCTTCGTCCGCAAACGCTGGTGCAACTGAGAGT Majority
              6910      6920      6930      6940      6950
6771 TCACTGTATTAGTTGTTGCTTCATCCGCAAACGCTGGTGCAACTGAGAGC 2603_a12.seq
6775 TCACAGTATCAGTTGTTACGTCGTCCGCAAACGCTGGTGCAACTGAGAAT nem316_a12.seq

     AGTGACGTTAAGGTCAGTAGCAGTGTCGAGAACATTGTAAGATATTTGTT Majority
              6960      6970      6980      6990      7000
6821 AATGACGTTAAAGTCAGTAACAATGCCGAGAACATTGCAAAATATTTGTT 2603_a12.seq
6825 AGTGACGTTAAGATCAATAGCAGTGTCGAGAACACTGTAAGACATTTGTT nem316_a12.seq

     GATTTTTTTCATTTCTATCTCCTTCTTATTTTAGTTAATCAACATGGTTA Majority
              7010      7020      7030      7040      7050
6871 GATTCTTTTCATTTCTATCTCCTTCTTATTTTAGTTAATCAACATGATTA 2603_a12.seq
6875 GATTTTTTTCATTTCTATCTCCTTCTTATTCTAGTTAATCAACATAGATA nem316_a12.seq

     ATAATATGCGGATTTTAATATTACCGCAGCACCACTCCTTTCAAGTCATG Majority
              7060      7070      7080      7090      7100
6921 ATAATATGCGGATTTTAATAC-ACCGCAGCACCACTCCCTTCAAGTCATG 2603_a12.seq
6925 ATAATATACGGATTATAATATTACCGCAGCACCACTCCTTTC-------- nem316_a12.seq

     GAATTTTATTTAATTAATTAAGAATACTAAAGCGCATGATTTTTAATCTT Majority
              7110      7120      7130      7140      7150
6970 GAATTTAGTTAATTAATTAAGAATACTAAAGCGCATAATTTTTAATCTT 2603_a12.seq
6967 GAATTATATTTAATTAATTAAGAATACTAAACCACATGATTCTTAATCTT nem316_a12.seq
```

FIGURE 20J

```
        TTTTTCTGGATATATCACTAGATTTCTTATATCTTTTCCAAATATAAATT Majority
             7160      7170      7180      7190      7200
7020    TTTTGATGGACATATCACTAGATTTCTTATACCTTTTCCAAATATAAATT 2603_a12.seq
7017    TTTCTCTAGATATATCACTAGATTTCTTATATCTTTTCCAAATATAAATT nem316_a12.seq

        CCACCTGCAATAGACATCATAGCTCCACCTATTAAAATGAAAGATAGAAT Majority
             7210      7220      7230      7240      7250
7070    CCACCTGCAATAGACATCATAGCTCCACCTATTAAAATGAAAGATAGAAT 2603_a12.seq
7067    CCACCTGCAATAGACATCATAGATCCACCTATTAAAATGAAAGATAGAAT nem316_a12.seq

        TCCTTTCCCACCTGTCATCGGAATAATTCCTTTTGGTGGAATATGCGTGT Majority
             7260      7270      7280      7290      7300
7120    TCCTTTCCCACCTGTCATAGGAATAATTCCTTTTGGTGGAATATGCGTGT 2603_a12.seq
7117    TCCTTTCCCACCTGTCATCGGAATAATTCCTTTTGGTGGAATATGCGTGT nem316_a12.seq

        TGGTAATTAAATGCTTGTCACCTTCCTCATGATATTCAGAAATCTGTTTA Majority
             7310      7320      7330      7340      7350
7170    TGGTAATTAAATGCTTGTCACCTTCCTCATGATATTCAGAAATCTGTTTA 2603_a12.seq
7167    TGGTAATTAAATGCTTGTCACCTTCCTCATGATATTCAGAAATCTGTTTA nem316_a12.seq

        TTAACAGCTATTATATTTTTTATCGATCCTTTAACCACTTCAAAAGTTAA Majority
             7360      7370      7380      7390      7400
7220    TTAACAGCTATTATATTTTTTATCGATCCTTTAACCACTTCAAAAGTTAA 2603_a12.seq
7217    TTAACAGCTATTATATTTTGTATCGATCCTTTAACAACTTCAAAAGTTAA nem316_a12.seq

        AATTGGTTTATTAGTAATTTTTTGATAATCCTTCGGCGAAACTGCTTCTA Majority
             7410      7420      7430      7440      7450
7270    AATTGGTTTATTAGTAATTTTTTGATAATCCTCCGGCGAAACTGCTTCTA 2603_a12.seq
7267    AATTGGTTTATTAGTAATTTTTTGATAATCCTTCGGCGAAACTGCTTCTA nem316_a12.seq

        TTAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCGTTT Majority
             7460      7470      7480      7490      7500
7320    TTAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCGTTT 2603_a12.seq
7317    TTAACTGATATTTGCCATCTTTCAAATCTTTGTAAGAAATTTTGCCGTTT nem316_a12.seq

        TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGGTAAATAAAGTTT Majority
             7510      7520      7530      7540      7550
7370    TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGGTAAATAAAGTTT 2603_a12.seq
7367    TCTCCCGTCACTACTTTTGAATTATTATTTTTTATTGGTAAATAAAGTTT nem316_a12.seq

        ATAATCTTCATTAAATTCTTGAAGTTCAAACGTAGCTCCTTTGAGAAGCA Majority
             7560      7570      7580      7590      7600
7420    ATAATCTTCATTAAATTCTTGAAGTTCAAACGTAGCTCCTTTGAGAAGCA 2603_a12.seq
7417    ATAATCTTCATTAAATTCTTGAAGTTCAAACGTAGCTCCTTTGAGAAGCA nem316_a12.seq

        ACTTATTATTATCTTTATCAACTTTTGTAAATTCAATTTCACCTAACTTC Majority
             7610      7620      7630      7640      7650
7470    ACTTATTATTATCTTTATCAACTTTTATAAATTCAATTTCACCTAACTTC 2603_a12.seq
7467    ACTTATTATTATCTTTATCAACTTTTGTAAATTCAATTTCACCTAACTTC nem316_a12.seq

        TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACGAATTTT Majority
             7660      7670      7680      7690      7700
7520    TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACGAATTTT 2603_a12.seq
7517    TTCTCGTTTTTAATCGTTATTGTAGGATATTCTCTCACATCACGAATTTT nem316_a12.seq

        AGGGATTGGAAAATCTCTAAGTGTATTAGGATCCTCTGATTTAGGATTCA Majority
             7710      7720      7730      7740      7750
7570    AGGGATTGGAAAATCTCTAAGTGTATTAGGATCCTCTGACTTAGGATTCA 2603_a12.seq
7567    AGGGATTGGAAAATCTCTAAGTGTATTAGGATCCTCTGATTTAGGATTCA nem316_a12.seq

        ATGTTGTTCTACCATTAGTGTCATAGAATTTGTTACTTATAAAACTGTCA Majority
             7760      7770      7780      7790      7800
7620    ATGTTGTTCTACCATTAGTGTCATAGAATTTGTTACTTATAAAACTGTCA 2603_a12.seq
7617    ATGTTGTTCTACCATTAGTGTCATAGAATTTGTTACTTATAAAACTGTCA nem316_a12.seq
```

FIGURE 20K

TCTAGTTTCACATCATATGTGAGTGTTACTTTTTGTCCTTCTCCTAAGTT Majority
7810 7820 7830 7840 7850
7670 TCTAGTTTCACATCATATGTGAGTGTTACTTTTTGACCTTCTCCTAAATT 2603_a12.seq
7667 TCTAGTTTCACATCATATGTGAGTGTTACTTTTTGTCCCTCCCCTAAGTT nem316_a12.seq

CAAACCTCTAACGTAGAGTTTATTTTTGATGTATTCTAATTTAACCCCTT Majority
7860 7870 7880 7890 7900
7720 CAAACCTCTAACATAGAGTTTATTTCCGATGTATTCTAATTTAACCCCCT 2603_a12.seq
7717 CAAACCTCTAACGTAGAGTTTATTTTTGATGTATTCTAATTTAACCCCTT nem316_a12.seq

TAAGTATTCCACCATCATTATTAGGCCCACCAGTTGCAATGCTATCTTTC Majority
7910 7920 7930 7940 7950
7770 TAAGTATTCCACCATCATTATTAGGCCCACCAGTTGCAATACCATCCTTC 2603_a12.seq
7767 TAAGTATTCCACCATCATTATTAGGCCCACCAGTTGCAATGCTATCTTTC nem316_a12.seq

ATTATACTTCCATCATTTCCCTGTAAAGTATAATCACTTGGTTGTAATGT Majority
7960 7970 7980 7990 8000
7820 ATTAGACTTCCATCATTTCCCTGTAAAGTATAATCACTTGGCTGTAATGT 2603_a12.seq
7817 ATTATACTTCCATCATTTCCCTGTAAAGTATAATCACTTGGTTGCAATGT nem316_a12.seq

TTGTCCGTTGCCAAGCTGTAAATTGATTTTGTCACCCATAGGATCTTCTA Majority
8010 8020 8030 8040 8050
7870 TTGTCCATTACCAAGCTGTAAATTGATTTTATCACCCATAGGATCTTCGA 2603_a12.seq
7867 TTGTCCGTTGCCAAGCTGTAAATTGATTTTGTCACCCATAGGATCTTCTA nem316_a12.seq

TAGTTCCATTAACAATTGAGTTTTCTTTTGTTAAAATCTTTTCAAATTGT Majority
8060 8070 8080 8090 8100
7920 TAGTTCCATTAACAATTGAGTTTTCTTTTGTTAAAATCGTTTCAAATTGT 2603_a12.seq
7917 TAGTTCCATTAACAATTGAGTTTTCTTTTGTTAAAACCTTTTCAAATTGT nem316_a12.seq

TGCTGAATTTTAGATAAAATTTCATTGTTAGATGTATCGGCTGAAGTTAC Majority
8110 8120 8130 8140 8150
7970 TGCTGAATTTTAGATAAAATTTCATTGTTAGATGTATCGGCTGAAGTTAC 2603_a12.seq
7967 TGCTGAATTTTAGATAAAATTTCATTGTTAGATGCATCGGATGAAGTTAC nem316_a12.seq

TATCGGGGTGTAGTACTCAGGTTTGGAAGAGAATGACTTCATTAGTTCTG Majority
8160 8170 8180 8190 8200
8020 GATAGGGGTGTAGTACTCAGGTTTGGAAGAGAACGACCTCATTAGTTCTG 2603_a12.seq
8017 TATCGGGGTATAATACTCAGGTTTAGAAGAGAATGACTTCATTAGTTCTG nem316_a12.seq

TTATTTCTCCATCTGAAAGTTTAAAAGCTTCCTCTTTCAATTTTTGAAAA Majority
8210 8220 8230 8240 8250
8070 TGATTTCTCCATCTGAAAGTTTAAAAGCTTCCTCTTTCAATTTTTGAAAA 2603_a12.seq
8067 TTATTTCCCCATCTGAAAGTTGAAAAGCTTCCTCTTTCAATTTTTGAAAA nem316_a12.seq

GTACCATCTTGATTTTTCTTATACTCCTCATTATAAACTTGTCTAAAAGC Majority
8260 8270 8280 8290 8300
8120 GTACCATCTTGATTTTTCTTATACTCCTCATTATAAACTTGTCTAAAACC 2603_a12.seq
8117 GTACCATCTTGATTTTTCTTATAATCCTCATTATAAACTTGTCTAAAAGC nem316_a12.seq

AGATATATCTATACCAAAATTAAAGATGTCATAATTTTTCTGTTTTAAAC Majority
8310 8320 8330 8340 8350
8170 AGATATATCGATACCAAAATTAAAAATGTCATAATTTTTCTGTTTTAAAC 2603_a12.seq
8167 AGATATATCTATACCAAAATTAAAGATGTCATAATTTTTCTGTTTTAAAC nem316_a12.seq

TATTTATATAAAGTTTGGTTGGTGTTCCATGTTCTTTTACTGGTCCATTT Majority
8360 8370 8380 8390 8400
8220 TATTTATATAAAGTTTGGTTGGTGTTCCATGTTCTTTCACTGGTCCATTT 2603_a12.seq
8217 TATTTATATAAAGTTTGGTTGGTGTTCCATGTTCTCTTACTGGTCCATTT nem316_a12.seq

CGATAAATTGTACCTTTAGGGTAATTAAGATTTAAATCTAAATAATGAAG Majority
8410 8420 8430 8440 8450
8270 CGATAAATTGTACCTTTAGGGTAATTAAGATTTAAATCTAAATAATGAAG 2603_a12.seq
8267 CGATAAATTGTACCTTTAGGGTAATTAAGATTTAAATCTAAATAATGAAG nem316_a12.seq

FIGURE 20L

TTTTTGTAAGTTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA Majority

8460 8470 8480 8490 8500

8320 TTTTTGTAAGTTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA 2603_a12.seq
8317 TTTTTGTAAGTTTCCAGAGATTATCTGTGTTTGATAACTATCTAAGGGAA nem316_a12.seq

ACAAAAAGTAACTCTCCCCATTTCCTTTTATATCCTCGGGCTTATCAGTA Majority

8510 8520 8530 8540 8550

8370 ACAAAAAGTAACTCTCCCCATTTCCTTTTATATCCTCGGGCTTATCAGTA 2603_a12.seq
8367 ACAAAAAGTAACTCTCCCCATTTCCTTTTATATCCTCGGGCTTATCAGTA nem316_a12.seq

AGTAGAAAATTACTTTTATTTAGATATCCATTTTTTTTCATTTGTTCAAA Majority

8560 8570 8580 8590 8600

8420 AGTAGAAAATTACTTTTATTTAGATATCCATTTTTTTTCATTTGTTCAAA 2603_a12.seq
8417 AGTAGAAAATTACTTTTATTTAGATATCCATTTTTTTTCATTTGTTCAAA nem316_a12.seq

TTGGCTTTCATATGATGCACCCAGTTTAAAATTATTAATAGCATATGATC Majority

8610 8620 8630 8640 8650

8470 TTGGCTTTCATATGATGCACCCAGTTTAAAATTATTAATAGCATATGATC 2603_a12.seq
8467 TTGGCTTTCATATGATGCACCCAATTTAAAATTATTAATAGCATATGATC nem316_a12.seq

TTGTTGGAACACCATCAGTTATATGAACAATAATTTTTTGACTATTTCGA Majority

8660 8670 8680 8690 8700

8520 TCGTAGGAACACCATCAGTTACATGAACAATAATTTTTTGACTATTTCGA 2603_a12.seq
8517 TTGTTGGAACACCATCAGTTATATGAACAATAATTTTTTGACTATTTCGA nem316_a12.seq

TTTACTTGACTCAAAATATCATCTGCCTCCATGAAGGCTTTCATAGTAAA Majority

8710 8720 8730 8740 8750

8570 TTTACTTGACTCAAAATATCATCTGCCTCCATGAAGGCTTTCATAGTAAA 2603_a12.seq
8567 TCTACTTGACTCAAAATATCATCTGCCTCCATGAAGGCTTTCATAGTAAA nem316_a12.seq

TGTTTCTCCTACTTTACTAAGATAGTACTGCTTTTGTTGCTCTGGAGTTA Majority

8760 8770 8780 8790 8800

8620 TGTTTCTCCTACTTTACTAAGATAGTACTCCTTTTGTTGCTCTGGAGTTA 2603_a12.seq
8617 TGTTTCCCCTACTTTACTAAGATAGTACTGCTTTTGTTGCTCTGGAGTAA nem316_a12.seq

GTCCGTTTGTAGTTGATCCCCATTTAGCTTTAGGAGCTTCTGTCGGAATC Majority

8810 8820 8830 8840 8850

8670 ATCCATTGGTAGTAGATCCCCACTTAGCTTTAGGAGCTTCTGTCGGAATC 2603_a12.seq
8667 GTCCGTTTGTAGTTGATCCCCATTTAGCTCTAGGAGCTTCTGTAGGAATC nem316_a12.seq

CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT Majority

8860 8870 8880 8890 8900

8720 CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT 2603_a12.seq
8717 CTTTTTATAATCTCTTCAGCATTATTTGTTAATTGTTTATGACTATAATT nem316_a12.seq

CTCTGTCTGAATTGTGAACTTAGTTTGAAGGCCATAATATTTATCATCTT Majority

8910 8920 8930 8940 8950

8770 CTCTGTCTGAATTGTGAACTTAGTTTGAAGGCCATAATATTTATCATCTT 2603_a12.seq
8767 CTCTGTCTGAATTGTGAACTTAGTTTGAAGGCCATAATATTTATCATCTT nem316_a12.seq

CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA Majority

8960 8970 8980 8990 9000

8820 CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA 2603_a12.seq
8817 CTTTAAATCCTTTTACGACATCTACACTCCTACCATCAAAAATATCTGAA nem316_a12.seq

CCATAGGTAACTAATGCAACCCTATTATCACTGTTTGCTCCTAAAATATC Majority

9010 9020 9030 9040 9050

8870 CCATAGGTAACTAATGCAACCCTATTATCACTGTTTGCTCCTAAAATATC 2603_a12.seq
8867 CCATAGGTAACTAATGCAACCCTATTATCACTGTTTGCTCCTAAAATATC nem316_a12.seq

TTTTACTGCGGTCCCAAGAGCTTCGGCAGCTTTCTTGGCTTTATTATGCC Majority

9060 9070 9080 9090 9100

8920 TTTTACTGCGGTCCCAAGAGCTTCGGCAGCTTTCTTGGCTTTATTATGCC 2603_a12.seq
8917 TTTTACTGCGGTCCCAAGAGCTTCGGCAGCTTTCTTGGCTTTATTATGCC nem316_a12.seq

FIGURE 20M

```
     TTTGAAAATTTGGCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG Majority
          9110      9120      9130      9140      9150
8970 TTTGAAAATTTGGCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG 2603_a12.seq
8967 TTTGAAAATTTGGCCATCGTTATTCATTGAGTTAGAATTATCGAGTACG nem316_a12.seq

     AAGACAACATCTAACGGCTTTTGTTTGTCCACTGGTTTTACTATGGTTTT Majority
          9160      9170      9180      9190      9200
9020 AAGACAACATCTAACGGCTTTTGTTTGTCCACTGGTTTTACTATGGTTTT 2603_a12.seq
9017 AAGACAACATCTAACGGCTTTTGTTTGTCCACTGGTTTTACTATGGTTTT nem316_a12.seq

     TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA Majority
          9210      9220      9230      9240      9250
9070 TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA 2603_a12.seq
9067 TCCACTGACAGTTAACTCAATTTTATATTTATTATGAGCTAAATCACCTA nem316_a12.seq

     CTTCTGAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC Majority
          9260      9270      9280      9290      9300
9120 CTTCTGAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC 2603_a12.seq
9117 CTTCTGAAATACGTTTAGATAATGTTCCCTCTGGAATTTCTCTTATATGC nem316_a12.seq

     TCACCTTCACTTGAATATGGGTTAACTGCTTTTGCCTCTGACTTTCCATT Majority
          9310      9320      9330      9340      9350
9170 TCACCTTCACTTGAATATGGGTTAACTGCTTTTGCCTCTGACTTTCCATT 2603_a12.seq
9167 TCACCTTCACTTGAATATGGGTTAACTGCTTTTGCCTCTGACTTTCCATT nem316_a12.seq

     TGGAACTGAACCTTTAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT Majority
          9360      9370      9380      9390      9400
9220 TGGAACTGAACCTTTAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT 2603_a12.seq
9217 TGGAACTGAACCTTTAACATGCTCAAGTTTATAAGATTCCTTTGTATCTT nem316_a12.seq

     CATAAATTCCTGTGGGGGGATACTGCTTATCTAGTTCTTCGTGATTTTGT Majority
          9410      9420      9430      9440      9450
9270 CATAAATTCCTGTGGGGGGATACTGCTTATCTAGTTCTTCCTGATTTTGT 2603_a12.seq
9267 CATAAATTCCTGTGGGGGGATACTGCTTATCTAGTTCTTCGTGATTTTGT nem316_a12.seq

     CCAATTGTGGAATTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT Majority
          9460      9470      9480      9490      9500
9320 CCAATTGTGGAATTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT 2603_a12.seq
9317 CCAATTGTGGAATTTTTATCACCACTATTTTGTATCGTAGTTTTTCCATT nem316_a12.seq

     ACTCTCAACCTTAACTTGCCAAGTCTGGTTAGTCTTTTTATAACCTTCGG Majority
          9510      9520      9530      9540      9550
9370 ACTCTCAACCTTAACTTGCCAAGTCTGGTTAGTCTTTTTATAACCTTCGG 2603_a12.seq
9367 ACTCTCAACCTTAACTTGCCAAGTCTGGTTAGTCTTTTTATAACCTTCGG nem316_a12.seq

     GCGCTGTTTCTTCTGATAAAGTATAATCTCCAGGTATGAGATTATCAAAA Majority
          9560      9570      9580      9590      9600
9420 GCGCTGTTTCTTCTGATAAAGTATAATCTCCAGGTATGAGATTATCAAAA 2603_a12.seq
9417 GCGCTGTTTCTTCTGATAAAGTATAATCTCCAGGTATGAGATTATCAAAA nem316_a12.seq

     GTAGCTTCACCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTGG Majority
          9610      9620      9630      9640      9650
9470 GTAGCTTCACCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTGG 2603_a12.seq
9467 GTAGCTTCACCTGTTAGCTCAGCAGTTACTTTTTCTATTTTACTTTCTGG nem316_a12.seq

     ATGAGCAGTAGTTTTTAAAACAAAGGTAGCTTTTGAAAGTGGTTTGTTCT Majority
          9660      9670      9680      9690      9700
9520 ATGAGCAGTAGTTTTTAAAACAAAGGTAGCTTTTGAAAGTGGTTTGTTCT 2603_a12.seq
9517 ATGAGCAGTAGTTTTTAAAACAAAGGTAGCTTTTGAAAGTGGTTTGTTCT nem316_a12.seq

     GGTCATCTGTCTTTTTAACAACTAACTTTCCTTTAGCACCATTTTCCGGT Majority
          9710      9720      9730      9740      9750
9570 GGTCATCTGTCTTTTTAACAACTAACTTTCCTTTAGCACCATTTTCCGGT 2603_a12.seq
9567 GGTCATCTGTCTTTTTAACAACTAACTTTCCTTTAGCACCATTTTCCGGT nem316_a12.seq
```

FIGURE 20N

ACGGTACTTTCCCCTAAAACATTGGTATTAAGCGGTATTTGCGACAAACA Majority

9760 9770 9780 9790 9800

9620 ACGGTACTTTCCCCTAAAACATTGGTATTAAGCGGTATTTGCGACAAACA 2603_al2.seq

9617 ACGGTACTTTCCCCTAAAACATTGGTATTAAGCGGTATTTGCGACAAACA nem316_al2.seq

AAAAAGACTTAACGTCAATATTTTAGAAAATTTTTGGTATTTTCTCATTT Majority

9810 9820 9830 9840 9850

9670 AAAAAGACTTAACGTCAATATTTTAGAAAATTTTTGGTATTTTCTCATTT 2603_al2.seq

9667 AAAAAGACTTAACGTCAATATTTTAGAAAATTTTTGGTATTTTCTCATTT nem316_al2.seq

TACAACTCCTATTGTGCCGAAATGTCGTTTCTAAATCTAAGATCAGATAC Majority

9860 9870 9880 9890 9900

9720 TACAACTCCTATTGTGCCGAAATGTCGTTTCTAAATCTAAGATCAGATAC 2603_al2.seq

9717 TACAACTCCTATTGTGCCGAAATGTCGTTTCTAAATCTAAGATCAGATAC nem316_al2.seq

AGAATATCCTAGAATATACAAACTATCACTTATTATGATATCAATAATTT Majority

9910 9920 9930 9940 9950

9770 AGAATATCCTAGAATATACAAACTATCACTTATTATGATATCAATAATTT 2603_al2.seq

9767 AGAATATCCTAGAATATACAAACTATCACTTATTATGATATCAATAATTT nem316_al2.seq

CTTATTATAAGGTATGGAATTTTAATGTTTTTTCCCAATTTTTGAATGAT Majority

9960 9970 9980 9990 10000

9820 CTTATTATAAGGTATGGAATTTTAATGTTTTTTCCCAATTTTTGAATAAT 2603_al2.seq

9817 CTTATTATAAGGTATGGAATTTTAAAGTTTTTTCCCAATTTTTGAATGAT nem316_al2.seq

TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATATTTCA Majority

10010 10020 10030 10040 10050

9870 TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATATTTCA 2603_al2.seq

9867 TTTTCTTTTTATTTGATAATCTTATTTTTTATTATCTTAGAAATACTTCA nem316_al2.seq

ATTAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTAAAATATTGCTT Majority

10060 10070 10080 10090 10100

9920 ATGAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTAAAATATTGCTT 2603_al2.seq

9917 ATTAGCTTAAGTAGTTGATTTTTCTTTTTTTATGTTTTAAAATATTGCTT nem316_al2.seq

AAAAATAATGTTTGAGAGAGAGTTTACTGAATTGATTGAAAATTATTTAG Majority

10110 10120 10130 10140 10150

9970 AAAAATAATGTTTGAGAGA--GTTTACTGAATTGATTGAAAATTATTTAG 2603_al2.seq

9967 AAAAATAATGTTTGAGAGAGAGTTTACTGAATTGATTGAAAATTATTTAG nem316_al2.seq

AAAAAGACATCCTTAATCAAATAAAACTTCTAACTTTATGCTATGATTAC Majority

10160 10170 10180 10190 10200

10018 AAAAAGACATCCTTAATCAAATAAAACTTCTAACTTTATGCTATGATTAC 2603_al2.seq

10017 AAAAAGACATCCTTAATCAAATAAAACTTCTAACTTTATGCTATGATTAC nem316_al2.seq

TACCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA Majority

10210 10220 10230 10240 10250

10068 TACCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA 2603_al2.seq

10067 TACCCTTCCATTACTCTAGACAAATCATGTCATCAACTTGGTTTATCTGA nem316_al2.seq

ACTACTTATTAGGAAATATTGTCATGATTTAACAACTTTATTTAACAGTC Majority

10260 10270 10280 10290 10300

10118 ACTACTTATTAGGAAATATTGTCATGATTTAACAACTTTATTTAACAGTC 2603_al2.seq

10117 ACTACTTATTAGGAAATATTGTCATGATTTAACAACTTTATTTAACAGTC nem316_al2.seq

AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT Majority

10310 10320 10330 10340 10350

10168 AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT 2603_al2.seq

10167 AACTCTCTCTGAATATCGAAAAGAGTACAATAGTCTACCAATCTAATGGT nem316_al2.seq

GTAACTAGAGAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT Majority

10360 10370 10380 10390 10400

10218 GTAACTAGAGAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT 2603_al2.seq

10217 GTAACTAGAGAACAAGCTTTCAAATATATTTATCATCAATCACACGTTTT nem316_al2.seq

FIGURE 20O

```
      ACAACTTTTAAAATTTTTGATCACGAATGATTCAGGCAGGTTACCTTTAA Majority
               10410     10420     10430     10440     10450
10268 ACAACTTTTAAAATTTTTGATCACGAATGATTCAGGCAGGTTACCTTTAA 2603_a12.seq
10267 ACAACTTTTAAAATTTTTGATCACGAATGATTCAGGCAGGTTACCTTTAA nem316_a12.seq

      CTTACTTTAGTGAAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCATA Majority
               10460     10470     10480     10490     10500
10318 CTTACTTTAGTGAAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCATA 2603_a12.seq
10317 CTTACTTTAGTGAAAAAATTTGGACTATCTTGTGCAACTGCTTATCGCATA nem316_a12.seq

      CGAAAACATATTAGTCCGTTACTAGAAAAACTTGGATTTCAGATTTTCAA Majority
               10510     10520     10530     10540     10550
10368 CGAAAACATATTAGTCCATTACTAGAAAAACTTGGATTTCAGATTGTCAA 2603_a12.seq
10367 CGAAAACATATTAGTCCGTTACTAGAAAAACTTGGATTTAAGATTTTCAA nem316_a12.seq

      AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCATTTT Majority
               10560     10570     10580     10590     10600
10418 AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCATTTT 2603_a12.seq
10417 AAATACTATTACCGGTGACGAGTATCGAATTCGCTATTTAATCGCACATT nem316_a12.seq

      TAAATGCTCGATTTGGTATAGAAGTTTATCCCTTGTCTAAGATGGATAAA Majority
               10610     10620     10630     10640     10650
10468 TAAATGCTCAATTTGGTATAGAAGTTTATCCCATGTCTAAGATGGATAAA 2603_a12.seq
10467 TAAATGCTCGATTTGGTATAGAAGTTTATCCCTTATCTAAGATGGATAAA nem316_a12.seq

      TTGCTTATCAAACGATTGTTATTAGAATACTCAACTACTTTTACTGCTTC Majority
               10660     10670     10680     10690     10700
10518 TTGCTCATCAAACGATTGTTATTAGAACACTCAACTACTTTTACTGCTTC 2603_a12.seq
10517 TTGCTTATCAAACGATTGTTATTAGAATACTCAACTACTTTTACTGCTTC nem316_a12.seq

      TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTGTTGTCTCTAT Majority
               10710     10720     10730     10740     10750
10568 TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTGTTGTCTCTAT 2603_a12.seq
10567 TCATTACTTCCCAAATACATTTATTTTCTTTGATACATTATTGTCTCTAT nem316_a12.seq

      CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCATCCCTTTTC Majority
               10760     10770     10780     10790     10800
10618 CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCATCCCTTTTC 2603_a12.seq
10617 CATGGAAACGTATTAATTATAATGTAGTTGTCCCTTACTCACCCCTTTTC nem316_a12.seq

      ATTGAACTACAAAATATATTTATCTATGATACATTACAATATTGTGTCAA Majority
               10810     10820     10830     10840     10850
10668 ACTGAACTACAAAATATATTTATCTATGATACATTACAATATTGTGTCAA 2603_a12.seq
10667 ATTGAACTACAAAATATATTTATCTATGATACATTACAATATTGTGTCAA nem316_a12.seq

      AAATGTTATTATAGATTCCTTTAAAATTAATTTAAAAAAAGACGATATAG Majority
               10860     10870     10880     10890     10900
10718 AAATGTTATTATAGATTCCTTTAAAATTAATTTAAAAAAAGACGATATAG 2603_a12.seq
10717 AAATGTTATTATAGATTCCTTTAAAATTAATTTAAAAAAAGACGATATAG nem316_a12.seq

      ACTATATTTTTCTTGCTTACCTTACTTCCCATAACTCTTTTTCTAATCCA Majority
               10910     10920     10930     10940     10950
10768 ACTATATTTTTCTTGCTTACCTTACTTCCCATAACTCTTTTTCTAATCCA 2603_a12.seq
10767 ACTATATTTTTCTTGCTTACCTTACTTCCCATAACTCTTTTTCTAATCCA nem316_a12.seq

      AATTGGACTGAGAAGCGTATCGATAATGTAATAGCTATTTTCGAAAATTA Majority
               10960     10970     10980     10990     11000
10818 AATTGGACTGAGAAGCGTATCGATAATGTAATAGCTATTTTCGAAAATTA 2603_a12.seq
10817 AATTGGACTGAGAAGCGTATCGATAATGTAATAGCTATTTTCGAAAATTA nem316_a12.seq

      TCCCAAATTCCAAAAATTATTACAGCCACTCAAAGATGCTCTTCCCTTAT Majority
               11010     11020     11030     11040     11050
10868 TCCCAAATTCCAAAAATTATTACAGCCACTCAAAGATGCTCTTCCCTTAT 2603_a12.seq
10867 TCCCAAATTCCAAAAATTATTACAGCCACTCAAAGATGCCCTTCCCTTAT nem316_a12.seq
```

FIGURE 20P

```
      CTGGCTCCTATCATGATGAGTTGGTAAAAGTTGCTATCATTTTTTTCCGA Majority
          11060     11070     11080     11090     11100
10918 CTGGCTCCTATCATGATGAGTTGGTAAAAGTTGCTATC-TTTTTTTCCGA 2603_a12.seq
10917 CTGACTCCTATCATGACGAGTTGGTAAAAGTTGCTATCATTTTTTCCCGA nem316_a12.seq

      ACATTTATTTTAGGATTAAATCAATTAATCCCTGAAACAATTTCATTTCC Majority
          11110     11120     11130     11140     11150
10967 ACATTTATTTTAGGATTAAATCAATTAATCCCTGAAACAATTTCATTTCC 2603_a12.seq
10967 ACATTTATTTTAGGATTAAATCAATTAATCCCTGAAACAATTTCATTTCC nem316_a12.seq

      TTCATGGAACTATCATAGACATGATAAATTAACTACTATTCTCCGACCGA Majority
          11160     11170     11180     11190     11200
11017 TTCATGGAACTATCATAGACATGATAAATTAACTACTATTCTCCGACCGA 2603_a12.seq
11017 TTCATGGAACTATCATAGACATGATAAATTAACTACTATTCTCCGACCGA nem316_a12.seq

      TAATTACAAATTGGTTAAGTGAAATTGGAGAATACACGTTTAAGGAACAA Majority
          11210     11220     11230     11240     11250
11067 TAATTACAAATTGGTTAAGTAAAATTGGAGAATACACGTTTAAAGAACAA 2603_a12.seq
11067 TAATTACAAATTGGTTAAGTGAAATTGGAGAATACACGTTTAAGGAACAA nem316_a12.seq

      CATTTTCTTCTCCTTTGTGCTCATCTAGAAAGAATTATCAAAAATCATAT Majority
          11260     11270     11280     11290     11300
11117 CATTTTCTTCTCCTTTGTACTCATCTAGAAAGAATCATCAAAAATCATAT 2603_a12.seq
11117 CATTTTCTTCTCCTTTGTGCTCACCTAGAAAGAATTATCAAAAATCACAT nem316_a12.seq

      TCCTCCGATACAGATAGCCGTACTAACTACAGACTTTATTAATAACCAAA Majority
          11310     11320     11330     11340     11350
11167 TCCTCCGATACAGATAGCCGTACTAACTACAGACTTTATTAATAACCAAA 2603_a12.seq
11167 TCCTCCGATACAAATAGCCGTACTAACTACAGACTTTATTAATAACAAAA nem316_a12.seq

      TTTTAACAGAATGTTTATTACAGAGGTTTTCTTCTAAACAGATTCATTTC Majority
          11360     11370     11380     11390     11400
11217 TTTTAACAGAATGCTTATTACAGAGATTTTCTTCTAAAAAGATTCATTTC 2603_a12.seq
11217 TTTTAACAGAATGTTTATCACAGAGGTTTTCTTCTAAACAGATTCATTTC nem316_a12.seq

      CACCCTTACTATTTATTAACTGATGATCTTTCCAATATTACTAATCTTAA Majority
          11410     11420     11430     11440     11450
11267 CACCCTTACTATTTATTAACTGATGATCTTTCCAATATTACTAATCTTAA 2603_a12.seq
11267 CACCCTTACTATTTATTAACTGATGATCTTTCCAATATTACTAATCTTAA nem316_a12.seq

      TCCAGATATTATTATTACCAATCCAAAGCTTTCTTCCTTTATCAAACATG Majority
          11460     11470     11480     11490     11500
11317 TCCAGACATTATTATTACCAATAAAAAGCTTTCTCCCTTTATCAAACATG 2603_a12.seq
11317 TCCAGATATTATTATTACCAACCCAAAGCTTTCTTCCTTTATAAAACATG nem316_a12.seq

      AGATTTCTTCAGAGAGTTTAATTACATATATTGATTTGATTCATACTTCA Majority
          11510     11520     11530     11540     11550
11367 AGATTTCTTCAGAGAGTTTAATTACATATATTGATCTGATTAATACTTCA 2603_a12.seq
11367 AGATTTCTTCAGAGAGTTTAATTACATATATTGATTTGATTCATACTCCA nem316_a12.seq

      GACCAGATCAATCAAATCCAAGAAATTATTTCATCAATACAGGAAGAAAA Majority
          11560     11570     11580     11590     11600
11417 GACCAAATCAATCAAATCCAAAAAATTATTTCATCAATACAGGAAGAAAA 2603_a12.seq
11417 GACCAGATCAATCAAATCCAAGAAATTATTTCATCAATACAAGAAGAAAA nem316_a12.seq

      ATATTGTAAACTTTTTGCAAAAACTAATGAAATAACTACTCGTAGCTCCT Majority
          11610     11620     11630     11640     11650
11467 ATATTGCAAACTTTTTGCAAAAACTAATGAAATAACTAATCGTAGCTCCT 2603_a12.seq
11467 ATATTGTAAACTTTTTGCAAAAAACAATGAAATAACTACCCGTAGCTCCT nem316_a12.seq

      ATAACTCTTAAAAATTAACATTAAAAAGCTAGAGCATTGTGTAATGCTCT Majority
          11660     11670     11680     11690     11700
11517 ATAACTCTTAAAAATTAACATTAAAAAGCTAGAGCATTGTGTAATGCTCT 2603_a12.seq
11517 ATAACTCTTAAAAATTAACATTAAAAAGCTAGAGCATTGTGTAATGCTCT nem316_a12.seq
```

FIGURE 20Q

```
      AGCTTTTTAATGTTAATTTTTTTGAATAATATAATCCAACTTTTCAACTG Majority
            11710     11720     11730     11740     11750
11567 AGCTTTTTAATGTTAATTTTTTTGAATAATATAATCCAACTTTTCAACTG 2603_a12.seq
11567 AGCTTTTTAATGTTAATTTTTTTGAATAATATAATCCAACTTTTCAACTG nem316_a12.seq

      TTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTGT Majority
            11760     11770     11780     11790     11800
11617 TTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTGT 2603_a12.seq
11617 TTTTTTCCCATGTGAAATGTTCTTTAATTCTTTTAGCAATATTCTGTTGT nem316_a12.seq

      AGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCATG Majority
            11810     11820     11830     11840     11850
11667 AGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCATG 2603_a12.seq
11667 AGTTTCTCTCTTAATGCCTTATCTTTTACTAATAAATCAAGAGATTCATG nem316_a12.seq

      GAGTGACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTATCAA Majority
            11860     11870     11880     11890     11900
11717 GAGTGACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTATCAA 2603_a12.seq
11717 GAGTGACTGAGTATTTTCTTCCATGATGATTCCTAACTCAGGGCTATCAA nem316_a12.seq

      TAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGCACTTGAAAGT Majority
            11910     11920     11930     11940     11950
11767 TAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGCACTTGAAAGT 2603_a12.seq
11767 TAACTTCAACTGTTCCACCGCGATCTGTTGCAATAATAGCACTTGAAAGT nem316_a12.seq

      AGACCAGCTTCTAAAATAGAGGTTGGTAATCCCTCTGGATACATTGAAGG Majority
            11960     11970     11980     11990     12000
11817 AGACCAGCTTCTAAAATAGAGGTTGGTAATCCCTCTGGATACATTGAAGG 2603_a12.seq
11817 AGACCAGCTTCTAAAATAGAGGTTGGTAATCCCTCTGGATACATTGAAGG nem316_a12.seq

      GTAAACAAAGATATCAGTCTGTGCCATTAAAGACATAGTCTGTTCAAAGT Majority
            12010     12020     12030     12040     12050
11867 GTAAACAAAGATATCAGTCTGTGCCATTAAAGACATAGTCTGTTCAAAGT 2603_a12.seq
11867 GTAAACAAAGATATCAGTCTGTGCCATTAAAGACATAGTCTGTTCAAAGT nem316_a12.seq

      TTAATTTCCCCAAAAAGTTAATCTGTTTGGACTGATATTTCTCTTTCAAA Majority
            12060     12070     12080     12090     12100
11917 TTAATTTCCCCAAAAAGTTAATCTGTTTGGACTGATATTTCTCTTTCAAA 2603_a12.seq
11917 TTAATTTCCCCAAAAAGTTAATCTGTTTGGACTGATATTTCTCTTTCAAA nem316_a12.seq

      TGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTAAATAAACATTTTCAGA Majority
            12110     12120     12130     12140     12150
11967 TGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTAAATAAACATTTTCAGA 2603_a12.seq
11967 TGTGCTAATTCAGGTCCGTCTCCTGCAATCTGTAAATAAACATTTTCAGA nem316_a12.seq

      GTACTGTGACATCGAAAATGCTTCTAAGAGCAATTCAATGCCTTTTTCTT Majority
            12160     12170     12180     12190     12200
12017 GTACTGTGACATCGAAAATGCTTCTAAGAGCAATTCAATGCCTTTTTCTT 2603_a12.seq
12017 GTACTGTGACATCGAAAATGCTTCTAAGAGCAATTCAATGCCTTTTTCTT nem316_a12.seq

      TAATAATTCTACCAGCATAAGTGATGAAAATATCATCAGCAGATTTTTCA Majority
            12210     12220     12230     12240     12250
12067 TAATAATTCTACCAGCATAAGTGATGAAAATATCATCAGCAGATTTTTCA 2603_a12.seq
12067 TAATAATTCTACCAGCATAAGTGATGAAAATATCATCAGCAGATTTTTCA nem316_a12.seq

      AGGTAAGCCGTACCAGCAAAATCAGAGCCTAGACTTTCAGATACCGAATT Majority
            12260     12270     12280     12290     12300
12117 AGGTAAGCCGTACCAGCAAAATCAGAGCCTAGACTTTCAGATACCGAATT 2603_a12.seq
12117 AGGTAAGCCGTACCAGCAAAATCAGAGCCTAGACTTTCAGATACCGAATT nem316_a12.seq

      ATAAATAACTCCTTTAGCTTCTATATTAAAATGTTTTAACCATTCAACGC Majority
            12310     12320     12330     12340     12350
12167 ATAAATAACTCCTTTAGCTTCTATATTAAAATGTTTTAACCATTCAACGC 2603_a12.seq
12167 ATAAATAACTCCTTTAGCTTCTATATTAAAATGTTTTAACCATTCAACGC nem316_a12.seq
```

FIGURE 20R

```
      TTCTCTTGGATACCGCATAAAAATCTGGACGATAATGCTTAACACGCGCT Majority
                12360     12370     12380     12390     12400
12217 TTCTCTTGGATACCGCATAAAAATCTGGACGATAATGCTTAACACGCGCT 2603_a12.seq
12217 TTCTCTTGGATACCGCATAAAAATCTGGACGATAATGCTTAACACGCGCT nem316_a12.seq

      GTGAGAAGATGTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT Majority
                12410     12420     12430     12440     12450
12267 GTGAGAAGATGTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT 2603_a12.seq
12267 GTGAGAAGATGTTCATAGATAGCTCCAAAGAAATCTAAAAAACGATTATT nem316_a12.seq

      GACAGAAAAATGACTTGACCCATGGTCTAAAACAATACTAGGTAAATGGT Majority
                12460     12470     12480     12490     12500
12317 GACAGAAAAATGACTTGACCCATGGTCTAAAACAATACTAGGTAAATGGT 2603_a12.seq
12317 GACAGAAAAATGACTTGACCCATGGTCTAAAACAATACTAGGTAAATGGT nem316_a12.seq

      GATTCTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGTGTA Majority
                12510     12520     12530     12540     12550
12367 GATTCTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGTGTA 2603_a12.seq
12367 GATTCTTTGCAAAAGATAGCCCTTCTAGCGTTGTTAACTGAAAACGTGTA nem316_a12.seq

      TTACAAATCACAAAATCAATATTTTCATCTGAAACATATTTCATCAGCGT Majority
                12560     12570     12580     12590     12600
12417 TTACAAATCACAAAATCAATATTTTCATCTGAAACATATTTCATCAGCGT 2603_a12.seq
12417 TTACAAATCACAAAATCAATATTTTCATCTGAAACATATTTCATCAGCGT nem316_a12.seq

      GTTGTATTCTCGATTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT Majority
                12610     12620     12630     12640     12650
12467 GTTGTATTCTCGATTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT 2603_a12.seq
12467 GTTGTATTCTCGATTTTTGTTAATAATAGGATAGCGCTGCTTGACAATAT nem316_a12.seq

      TTTTGGTCGGTAAACGGTAAATTTTTCTACCCTTGTCTTCATCTATAATC Majority
                12660     12670     12680     12690     12700
12517 TTTTGGTCGGTAAACGGTAAATTTTTCTACCCTTGTCTTCATCTATAATC 2603_a12.seq
12517 TTTTGGTCGGTAAACGGTAAATTTTTCTACCCTTGTCTTCATCTATAATC nem316_a12.seq

      GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT Majority
                12710     12720     12730     12740     12750
12567 GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT 2603_a12.seq
12567 GGTAAATCACCATGATTAGTTGTTACAATAACAACACGGTAGCCACGCTT nem316_a12.seq

      AACCAAATCTGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGA Majority
                12760     12770     12780     12790     12800
12617 AACCAAATCTGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGA 2603_a12.seq
12617 AACCAAATCTGCTGTCATTTTATCTGTATAACGTTCAATACCTCCGAGGA nem316_a12.seq

      AGGGTAGATAATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCC Majority
                12810     12820     12830     12840     12850
12667 AGGGTAGATAATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCC 2603_a12.seq
12667 AGGGTAGATAATATCCTGAGAAAACAGCAACTGTTTTTACCTTATTTTCC nem316_a12.seq

      ATATTTATCCACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG Majority
                12860     12870     12880     12890     12900
12717 ATATTTATCCACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG 2603_a12.seq
12717 ATATTTATCCACTTTCATCAATAAGCCATCTTTTAAGCCTTTAATCATAG nem316_a12.seq

      CAACTATTTTTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAAT Majority
                12910     12920     12930     12940     12950
12767 CAACTATTTTTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAAT 2603_a12.seq
12767 CAACTATTTTTTTGCTCTTTTGCTCTTCTGCTACCAACACTCGAACAAAT nem316_a12.seq

      TCATTTCGCATAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT Majority
                12960     12970     12980     12990     13000
12817 TCATTTCGCATAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT 2603_a12.seq
12817 TCATTTCGCATAAATACTAAATATTTGTATCGCTTCTTCTTACCATATTT nem316_a12.seq
```

FIGURE 20S

```
      TTTTATAATATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTG Majority
               13010     13020     13030     13040     13050
12867 TTTTATAATATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTG 2603_a12.seq
12867 TTTTATAATATAGATCGCATTGCGTATCATGTAATATTTTCGAAATGGTG nem316_a12.seq

      AATGATTCAATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAG Majority
               13060     13070     13080     13090     13100
12917 AATGATTCAATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAG 2603_a12.seq
12917 AATGATTCAATACATGAAAAACATGGCCAAATTTTTTAACTCGTGAAGAG nem316_a12.seq

      TGTCCAATTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA Majority
               13110     13120     13130     13140     13150
12967 TGTCCAATTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA 2603_a12.seq
12967 TGTCCAATTTCGTGTAACAGACCAATAAAATTAACCTGATAAGTCTTATA nem316_a12.seq

      TCCCATCTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAA Majority
               13160     13170     13180     13190     13200
13017 TCCCATCTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAA 2603_a12.seq
13017 TCCCATCTCTGACAGACGATAATTCATTTCAGAGTCAACAAAATCAATAA nem316_a12.seq

      ACATCTCTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATT Majority
               13210     13220     13230     13240     13250
13067 ACATCTCTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATT 2603_a12.seq
13067 ACATCTCTTCTGCAAAGCCAGATGTTTCTTCGAAAACGCTCGTTTTCATT nem316_a12.seq

      AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCTTG Majority
               13260     13270     13280     13290     13300
13117 AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCTTG 2603_a12.seq
13117 AAAGCAGCCGAAGTAATACACTCTTCAATTTCTTTATAGTCAAATTCTTG nem316_a12.seq

      CATCACTAAATTTTCACGGTTCATATCTTGATACAAACAAGATAACATAC Majority
               13310     13320     13330     13340     13350
13167 CATCACTAAATTTTCACGGTTCATATCTTGATACAAACAAGATAACATAC 2603_a12.seq
13167 CATCACTAAATTTTCACGGTTCATATCTTGATACAAACAAGATAACATAC nem316_a12.seq

      CGACCTTAGGTAAATGAAGGTAATTTTCATAATTATCTATCAAATCACCT Majority
               13360     13370     13380     13390     13400
13217 CGACCTTAGGTAAATGAAGGTAATTTTCATAATTATCTATCAAATCACCT 2603_a12.seq
13217 CGACCTTAGGTAAATGAAGGTAATTTTCATAATTATCTATCAAATCACCT nem316_a12.seq

      AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCTTGTGC Majority
               13410     13420     13430     13440     13450
13267 AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCTTGTGC 2603_a12.seq
13267 AGGACAACCGAATCTTGATCTAAAGTCAAGAACCAATCAAATTCTTGTGC nem316_a12.seq

      TACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATCCCTTTATTTT Majority
               13460     13470     13480     13490     13500
13317 TACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATCCCTTTATTTT 2603_a12.seq
13317 TACTGCAAATTGACCGATACAGTTCAAAGCATATGCAATCCCTTTATTTT nem316_a12.seq

      CTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAATCGGCTACT Majority
               13510     13520     13530     13540     13550
13367 CTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAATCGGCTACT 2603_a12.seq
13367 CTGTTAAATAATCAACAGTTAGGTGCCCCTCTTCATTATAATCGGCTACT nem316_a12.seq

      AATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTACGATATAGAT Majority
               13560     13570     13580     13590     13600
13417 AATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTACGATATAGAT 2603_a12.seq
13417 AATTGAGAAATTTCTTCCTTATTTTTCGAGCCATTATCTACGATATAGAT nem316_a12.seq

      GTGGCTTACTTGAGGATAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA Majority
               13610     13620     13630     13640     13650
13467 GTGGCTTACTTGAGGATAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA 2603_a12.seq
13467 GTGGCTTACTTGAGGATAAATTGCTCGAATGTTCTGATCTAAGCGTTCAA nem316_a12.seq
```

FIGURE 20T

```
      TATTGGGGTTAAAGGTGACAATACCCGCTAAATATTTCATGTTCTATGCT Majority
          13660     13670     13680     13690     13700
13517 TATTGGGGTTAAAGGTGACAATACCCGCTAAATATTTCATGTTCTATGCT 2603_a12.seq
13517 TATTGGGGTTAAAGGTGACAATACCCGCTAAATATTTCATGTTCTATGCT nem316_a12.seq

      CTTTTCTAAAATCTCTAAATAACTGAATGACTGGTGCTTTGGTTATAAAA Majority
          13710     13720     13730     13740     13750
13567 CTTTTCTAAAATCTCTAAATAACTGAATGACTGGTGCTTTGGTTATAAAA 2603_a12.seq
13567 CTTTTCTAAAATCTCTAAATAACTGAATGACTGGTGCTTTGGTTATAAAA nem316_a12.seq

      ACGATACCGACATAGATAGTTACTGCTACTAAACTTTGAATGACATAATT Majority
          13760     13770     13780     13790     13800
13617 ACGATACCGACATAGATAGTTACTGCTACTAAACTTTGAATGACATAATT 2603_a12.seq
13617 ACGATACCGACATAGATAGTTACTGCTACTAAACTTTGAATGACATAATT nem316_a12.seq

      TACCAATGATACTGACATTTGAGTATTGATATAATAGAGTACAGCTCCAC Majority
          13810     13820     13830     13840     13850
13667 TACCAATGATACTGACATTTGAGTATTGATATAATAGAGTACAGCTCCAC 2603_a12.seq
13667 TACCAATGATACTGACATTTGAGTATTGATATAATAGAGTACAGCTCCAC nem316_a12.seq

      TAAGAGTAGCAGCAATTAAATAGCGCAGCATTCCTCTTGTTAATTCTTTA Majority
          13860     13870     13880     13890     13900
13717 TAAGAGTAGCAGCAATTAAATAGCGCAGCATTCCTCTTGTTAATTCTTTA 2603_a12.seq
13717 TAAGAGTAGCAGCAATTAAATAGCGCAGCATTCCTCTTGTTAATTCTTTA nem316_a12.seq

      AAAGTAAATACATCTCTTAAAGAGATAGCTTGATATAGGGAGACAATAAA Majority
          13910     13920     13930     13940     13950
13767 AAAGTAAATACATCTCTTAAAGAGATAGCTTGATATAGGGAGACAATAAA 2603_a12.seq
13767 AAAGTAAATACATCTCTTAAAGAGATAGCTTGATATAGGGAGACAATAAA nem316_a12.seq

      TTCAGTAATAACTGTAGAGATAATAGCTCCCATAGCACCTAAAATTGGTA Majority
          13960     13970     13980     13990     14000
13817 TTCAGTAATAACTGTAGAGATAATAGCTCCCATAGCACCTAAAATTGGTA 2603_a12.seq
13817 TTCAGTAATAACTGTAGAGATAATAGCTCCCATAGCACCTAAAATTGGTA nem316_a12.seq

      TTAAAAGTATATTAAGCACAACATTTGCCACAAGTCCAATAACTGCAGAC Majority
          14010     14020     14030     14040     14050
13867 TTAAAAGTATATTAAGCACAACATTTGCCACAAGTCCAATAACTGCAGAC 2603_a12.seq
13867 TTAAAAGTATATTAAGCACAACATTTGCCACAAGTCCAATAACTGCAGAC nem316_a12.seq

      ATTGTGTAAGCTTTTGTACGTCTTGAAGCCAGTAGATACTGTGTCCCTAA Majority
          14060     14070     14080     14090     14100
13917 ATTGTGTAAGCTTTTGTACGTCTTGAAGCCAGTAGATACTGTGTCCCTAA 2603_a12.seq
13917 ATTGTGTAAGCTTTTGTACGTCTTGAAGCCAGTAGATACTGTGTCCCTAA nem316_a12.seq

      AGCGTTACCATAAGAAATGCAAATGATCATCAAA                 Majority
          14110     14120     14130
13967 AGCGTTACCATAAGAAATGCAAATGATCATCAAA                 2603_a12.seq
13967 AGCGTTACCATAAGAAATGCAAATGATCATCAAA                 nem316_a12.seq
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 20U

```
      TCCACATCGGTCCAATTAACATATGACGTGGCGCATCACCAGTAATTCGG Majority
               10        20        30        40        50
  1   TCCACATCGGTCCAATTAACATATGACGTGGCGCATCACCAGTAATTCGG coh1_al2.seq
  1   TCCACATCGGTCCAATTAACATATGACGTGGCGCATCACCAGTAATTCGG a909_al2.seq

      TGAATAACAATATGTTTTGGAATAATCTCCAGTTGGTCACAAATAATCGA Majority
               60        70        80        90       100
 51   TGAATAACAATATGTTTTGGAATAATCTCCAGTTGGTCACAAATAATCGA coh1_al2.seq
 51   TGAATAACAATATGTTTTGGAATAATCTCCAGTTGGTCACAAATAATCGA a909_al2.seq

      AATATAGTCTTCTTGACTTAACAAACGTAAACGACCTTCATGGTAATCTC Majority
              110       120       130       140       150
101   AATATAGTCTTCTTGACTTAACAAACGTAAACGACCTTCATGGTAATCTC coh1_al2.seq
101   AATATAGTCTTCTTGACTTAACAAACGTAAACGACCTTCATGGTAATCTC a909_al2.seq

      TCTGCATTCTTGTATTAGTCATAAGATGCAGAAGGTGTAATTTTATACCC Majority
              160       170       180       190       200
151   TCTGCATTCTTGTATTAATCATAAGATGCAGAAGGTGTAATTTTATACCC coh1_al2.seq
151   TCTGCATTCTTGTATTAGTCATAAGATGCAGAAGGTGTAATTTTATACCC a909_al2.seq

      TGAATATCATTATCCGTAACACATCGACGAACATTTTCCACCATCATATC Majority
              210       220       230       240       250
201   TGAATATCATTATCCGTAACACATCGACGAACATTTTCCACCATCATATC coh1_al2.seq
201   TGAATATCATTATCCGTAACACATCGACGAACATTTTCCACCATCATATC a909_al2.seq

      ATGTGTCTCCCCTGGGAGACCATTTATTAGGTGAGAAACGATTTCTACTT Majority
              260       270       280       290       300
251   ATGTGTCTCCCCTGGGAGACCATTTATTAGGTGAGAAACGATTTCTACTT coh1_al2.seq
251   ATGTGTCTCCCCTGGGAGACCATTTATTAGGTGAGAAACGATTTCTACTT a909_al2.seq

      TAGGAGCTAATTCTCGTATTCTCTTAACAGTTTTTTTGTAAAGGTCATAT Majority
              310       320       330       340       350
301   TAGGAGCTAATTCTCGTACTCTCTTAACAGTTTTTTTGTAAAGGTCATAT coh1_al2.seq
301   TAGGAGCTAATTCTCGTATTCTCTTAACAGTTTTTTTGTAAAGGTCATAT a909_al2.seq

      GAATGTGCTCTATTTATTAATGCAGAAGTTGCTTCATAAGTTGTCTGAAG Majority
              360       370       380       390       400
351   GAATGTGCTCTATTTATTAATGCAGAAGTTGCTTCATAAGTTGTCTGAAG coh1_al2.seq
351   GAATGTGCTCTATTTATTAATGCAGAAGTTGCTTCATAAGTTGTCTGAAG a909_al2.seq

      GCCTAATTCTAAAGTCACATGCATTCTTTCAGAAAGTTCAGCGAGATAGT Majority
              410       420       430       440       450
401   GCCTAATTCTAAAGTCACATGCATTCTTTCAGAAAGTTCAGCGAGATAGT coh1_al2.seq
401   GCCTAATTCTAAAGTCACATGCATTCTTTCAGAAAGTTCAGCGAGATAGT a909_al2.seq

      ATATAGTTTCATCAGGTAAGCAATCCGGCCTTGTTCCGATGTTGATCCCG Majority
              460       470       480       490       500
451   ATATAGTTTCATCAGGTAAGCAATCCGGCCTTGTTCCGATGTTGATCCCG coh1_al2.seq
451   ATATAGTTTCATCAGGTAAGCAATCCGGCCTTGTTCCGATGTTGATCCCG a909_al2.seq

      ATAACTCCTGGCTCATTAATAGCCTGTTCGTAACGCTCTTTAATTATCTC Majority
              510       520       530       540       550
501   ATAACTCCTGGCTCATTAATAGCCTGTTCGTAACGCTCTTTAATTATCTC coh1_al2.seq
501   ATAACTCCTGGCTCATTAATAGCCTGTTCGTAACGCTCTTTAATTATCTC a909_al2.seq

      TAACTTAGCATGGGTATTGGTAAAATTTTGAAAATAGACTAAGTATTTAT Majority
              560       570       580       590       600
551   TAACTTAGCATGGGTATTGGTAAAATTTTGAAAATAGACTAAGTATTTAT coh1_al2.seq
551   TAACTTAGCATGGGTATTGGTAAAATTTTGAAAATAGACTAAGTATTTAT a909_al2.seq

      TAACCTCAGGCCACTTTCTATGCATGAAATCAATTTCTTTATAGAATTGT Majority
              610       620       630       640       650
601   TAACCTCAGGCCACTTTCTATGCATGAAATCAATTTCTTTATAGAATTGT coh1_al2.seq
601   TAACCTCAGGCCACTTTCTATGCATGAAATCAATTTCTTTATAGAATTGT a909_al2.seq
```

Figure 21

```
          TCACGAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC  Majority
                   660       670       680       690       700
651       TCACGAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC  cohl_a12.seq
651       TCACGAATAGGAGCTTCTGGAGCAACTATAGCATCCCCTGAACCAGAAAC  a909_a12.seq

          TGTGCAAAAAGTGCACCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC  Majority
                   710       720       730       740       750
701       TGTGCAAAAAGTGCACCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC  cohl_a12.seq
701       TGTGCAAAAAGTGCACCCTCCTCTAGCAACTGTTCCATCTCTGTTAGGAC  a909_a12.seq

          AGTCAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT  Majority
                   760       770       780       790       800
751       AGTCAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT  cohl_a12.seq
751       AGTCAAAACCAGCATCTATAGGTAATTTAAATATTTTTTCTCCAAAGAGT  a909_a12.seq

          TCTCGATAATAATCATTAATCGCACGATAACGTTTTTTCATAGGATAATT  Majority
                   810       820       830       840       850
801       TCTCGATAATAATCATTAATCGCACGATAACGTTTTTTCATAGGATAATT  cohl_a12.seq
801       TCTCGATAATAATCATTAATCGCACGATAACGTTTTTTCATAGGATAATT  a909_a12.seq

          GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA  Majority
                   860       870       880       890       900
851       GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA  cohl_a12.seq
851       GTATCACAATTTTAACTAAAATAACCTCACTACTACAATAAAACTAAAAA  a909_a12.seq

          AGATTGGAACGTCAGTTAGTTCCAATCTTTTATTTACTTCACTTTCTTTA  Majority
                   910       920       930       940       950
901       AGATTGGAACGTCAGTTAGTCCCAATCTTTTATTTACTTCACTTTCTTTA  cohl_a12.seq
901       AGATTGGAACGTCAGTTAGTTCCAATCTTTTATTTACTTCACTTTCTTTA  a909_a12.seq

          ACCAATCCTTGGCTAAAAAGATATACGCAGTTAGATTCAAAATACCATAA  Majority
                   960       970       980       990      1000
951       ACCAATCCTTGGATAAAAAGATATACGCAGTTAGATTCAAAATACCATAA  cohl_a12.seq
951       ACCAATCCTTGGCTAAAAAGATATACGCAGTTAGATTCAAAATACCATAA  a909_a12.seq

          GCAAGTATAAAACCAGCTAAAACATCTGTCGGAAAATGAACCCCTAGGTA  Majority
                  1010      1020      1030      1040      1050
1001      GCAAGTATAAAACCAGCTAAAACATCTGTCGGAAAATGAACCCCTAGGTA  cohl_a12.seq
1001      GCAAGTATAAAACCAGCTAAAACATCTGTCGGAAAATGAACCCCTAGGTA  a909_a12.seq

          AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAATGTACCTTGGC  Majority
                  1060      1070      1080      1090      1100
1051      AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAATGTACCTTGGC  cohl_a12.seq
1051      AATACGAGATAACCCAATTAAAAAAATGAGCAAACCCAAAATACCTTGGC  a909_a12.seq

          ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA  Majority
                  1110      1120      1130      1140      1150
1101      ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA  cohl_a12.seq
1101      ACAACAGTTTCCATATACTCTTAGGCATATAGTACTGCAATAAAATAATA  a909_a12.seq

          CTACTCCCAAATATCATAAATGTTCCCATCGAGTGCCCACTGGGAAACGA  Majority
                  1160      1170      1180      1190      1200
1151      CTACTCCCAAATATCATAAATGTTCCCATCGAGTGCCCACTGGGAAACGA  cohl_a12.seq
1151      ATACTCCCAAATATCATAAATGTTCCCATCGAGTGCCCACTGGGAAACGA  a909_a12.seq

          ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA  Majority
                  1210      1220      1230      1240      1250
1201      ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA  cohl_a12.seq
1201      ATAGCCACCTGCAAATACTAAATGGGTTAAAGTTGGTCTCACTCTTTGAA  a909_a12.seq

          AAATAAGTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTTACT  Majority
                  1260      1270      1280      1290      1300
1251      AAATAAGTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTTACT  cohl_a12.seq
1251      AAATAAGTTTTAAAGAAAGTATACATATACCAGAGATAATAGCATTTACT  a909_a12.seq
```

FIGURE 21A

GCGATAAATCTAGCTTGAGGATACCACTTCTTAAGGTAACAGAAAGTGAC Majority
1310 1320 1330 1340 1350
1301 GCGATAAATCTAGCTTGAGGATACCACTTCTTAAGGTAACAGAAAGTGAC coh1_a12.seq
1301 GCGATAAATCTAGCTTGAGGATACCACTTCTTAAGGTAACAGAAAGTGAC a909_a12.seq

GCTCATAATCGCAATAGCTATCTGGCTTACAGTATTACCAACCACAGTGA Majority
1360 1370 1380 1390 1400
1351 GCTCATAATCGCAATAGCTATCTGGCTTACAGTATTACCAACCACAGTGA coh1_a12.seq
1351 GCTCATAATCGCAATAGCTATCTGGCTTACAGTATTACCAACCACAGTGA a909_a12.seq

TTAACTTGAAAAATCTTGTAGAAAGATTTGGCAACTGTCCTCTAACACTT Majority
1410 1420 1430 1440 1450
1401 TTAACTTAAAAAATCTTGTAGAAAGATTTGGCAACTGTCCTCTAACACTT coh1_a12.seq
1401 TTAACTTGAAAAATCTTGTAGAAAGATTTGGCAACTGTCCTCTAACACTT a909_a12.seq

TCTTGAATAGTTTGGTCAAATGCGATTACAGTGTCGGGCCAATATTTGAT Majority
1460 1470 1480 1490 1500
1451 TCTTGAATAGTTTGGTCAAATGCGATTACAGTGTCGGGCCAATATTTGAT coh1_a12.seq
1451 TCTTGAATAGTTTGGTCAAATGCGATTACAGTGTCGGGCCAATATTTGAT a909_a12.seq

GACCAATCCTAAACTGAAAAATAAGATAATAGCAATAAATGCTTGAATAA Majority
1510 1520 1530 1540 1550
1501 GACCAATCCTAAACTGAAAAATAAGATAATAGCAATAAATGCTTGAATAA coh1_a12.seq
1501 GACCAATCCTAAACTGAAAAATAAGATAATAGCAATAAATGCTTGAATAA a909_a12.seq

GTTTACTATTTTGACGAGATAAGATTAGTCTTTTTATATCTTTCTAATAT Majority
1560 1570 1580 1590 1600
1551 GTTTACTATTTTGACGAGATAACATTAGTCTTTTTATATCTTTCTAATAT coh1_a12.seq
1551 GTTTACTATTTTGACGAGATAACATTAGTCTTTTTATATCTTTCTAATAT a909_a12.seq

TGGCAAACAAGCCACGTAAGTTAGATAGAAAACAATCGAAATTAAAATTC Majority
1610 1620 1630 1640 1650
1601 TGGCAAACAAGCCACGTAAGTTAGATAGAAAACAATCGAAATTAAAATTC coh1_a12.seq
1601 TGGCAAACAAGCCACGTAAGTTAGATAGAAAACAATCGAAATTAAAATTC a909_a12.seq

CCTCAACGATATTAAATGGAATAACCATTGTTAAAAGGTAATTGCCTACA Majority
1660 1670 1680 1690 1700
1651 CCTCAACGATATTAAATGGAATAACCATTGTTAAAAGGTAATTGCCTACA coh1_a12.seq
1651 CCTCAACGATATTAAATGGAATAACCATTGTTAAAAGGTAATTGCCTACA a909_a12.seq

CCAATAAATGTTCTGATATCAAAGTTAGCAAATATAGCATACAAAGGAAT Majority
1710 1720 1730 1740 1750
1701 CCAATAAATGTTCTGATATCAAAGTTAGCAAATATAGCATACAAAGGAAT coh1_a12.seq
1701 CCAATAAATGTTCTGATATCAAAGTTAGCAAATATAGCATACAAAGGAAT a909_a12.seq

CGCAAAGACATAGTTGAGAGCTACCATAGATACAGTCAAGCTAACTGTAC Majority
1760 1770 1780 1790 1800
1751 CGCAAAGACATAGTTGAGAGCTACCATAGATACAGTCAAGCTAACTGTAC coh1_a12.seq
1751 CGCAAAGACATAGTTGAGAGCTACCATAGATACAGTCAAGCTAACTGTAC a909_a12.seq

CAAATAAACTAGCTTTAATAAAATCTTTTGCACTCTCTCTATTTTTCCAG Majority
1810 1820 1830 1840 1850
1801 CAAATAAACTAGCTTTAATAAAATCTTTTGCACTCTCTCTATTTTTCCAG coh1_a12.seq
1801 CAAATAAACTAGCTTTAATAAAATCTTTTGCACTCTCTCTATTTTTCCAG a909_a12.seq

AAAATAGCGAAACTTGCTAAAAATAGAGCTAGAGCAACCATATTCATCGG Majority
1860 1870 1880 1890 1900
1851 AAAATAGCGAAACTTGCTAAAAATAGAGCTAGAGCAACCATATTCATCGG coh1_a12.seq
1851 AAAATAGCGAAACTTGCTAAAAATAGAGCTAGAGCAACCATATTCATCGG a909_a12.seq

TAAACCGATAAAGGTTTCTGGACCACGATTAGCAAGTATAACTTTTAAAA Majority
1910 1920 1930 1940 1950
1901 TAAACCGATAAAGGTTTCTGGACCACGATTAGCAAGTATAACTTTTAAAA coh1_a12.seq
1901 TAAACCGATAAAGGTTTCTGGACCACGATTAGCAAGTATAACTTTTAAAA a909_a12.seq

FIGURE 21B

113/487

```
     GTGATCTTAATAAGAGTACACCATAACTTGATTTCAAATCAAATAAAATA Majority
              1960      1970      1980      1990      2000
1951 GTGATCTTAATAAGAGTACACCATAACTTGATTTCAAATCAAATAAAATA cohl_al2.seq
1951 GTGATCTTAATAAGAGTACACCATAACTTGATTTCAAATCAAATAAAATA a909_al2.seq

     AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAAATTCTGC Majority
              2010      2020      2030      2040      2050
2001 AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAAATTCTGC cohl_al2.seq
2001 AAAGCAACTAACATCGGAAGGATTGAAAAATCAACCTTTAAAAAATTCTGC a909_al2.seq

     TCCTGGTATTAATGGAAATGAAACCATCATCAATACAAAAGATAAGGCAG Majority
              2060      2070      2080      2090      2100
2051 TCCTGGTATTAATGGAAATGAAACCATCATCAATACAAAAGATAAGGCAG cohl_al2.seq
2051 TCCTGGTATTAATGGAAATGAAACCATCATCAATACAAAAGATAAGGCAG a909_al2.seq

     AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC Majority
              2110      2120      2130      2140      2150
2101 AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC cohl_al2.seq
2101 AAAGAATGGCGATTGTCACCATTTTACGTGTATTTGTCATAAAAAAATTC a909_al2.seq

     CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC Majority
              2160      2170      2180      2190      2200
2151 CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC cohl_al2.seq
2151 CTCCAATTTAAATAAATTGAAAGAAGCTCCAAAGGTAAGCGTAGGTACGC a909_al2.seq

     GAAAAAAACCTTTGTCTTCTCCCATCCAGACTTTACTGTCGGTTGTGGAA Majority
              2210      2220      2230      2240      2250
2201 GAAAAAAACCTTTGTCTTCTCCCATCCAGACTTTACTGTCGGTTGTGGAA cohl_al2.seq
2201 GAAAAAAACCTTTGTCTTCTCCCATCCAGACTTTACTGTCGGTTGTGGAA a909_al2.seq

     TCTCACCACATCAGCTTTCGCTCGCGGACTGATGCTTCACAACTGACAAA Majority
              2260      2270      2280      2290      2300
2251 TCTCACCACATCAGCTTTCGCTCGCGGACTGATGCTTCACAACTGACAAA cohl_al2.seq
2251 TCTCACCACATCAGCTTTCGCTCGCGGACTGATGCTTCACAACTGACAAA a909_al2.seq

     TAAGTTGGAAGCGATTACCGCCGGTCGGGAATTACACCCTGCCCTGAAGA Majority
              2310      2320      2330      2340      2350
2301 TAAGTTGGAAGCGATTACCGCCGGTCGGGAATTACACCCTGCCCTGAAGA cohl_al2.seq
2301 TAAGTTGGAAGCGATTACCGCCGGTCGGGAATTACACCCTGCCCTGAAGA a909_al2.seq

     CACCTATAGCATAACAAAAAAAAACTTGCAATTGCAAGTTTTTTAATCACT Majority
              2360      2370      2380      2390      2400
2351 CACCTATAGCATAACAAAAAAAAACTTGCAATTGCAAGTTTTTTAATCACT cohl_al2.seq
2351 CACCTATAGCATAACAAAAAAAAACTTGCAATTGCAAGTTTTTTAATCACT a909_al2.seq

     AATTAGTAGTAGATTGTATAATATTAATTTTTAACATCAATTAATTGACA Majority
              2410      2420      2430      2440      2450
2401 AATTAGTAGTAGATTGTATAATATTAATTTTTAACATCAATTAATTGACA cohl_al2.seq
2401 AATTAGTAGTAGATTGTATAATATTAATTTTTAACATCAATTAATTGACA a909_al2.seq

     GCGCACTAATACTCTAGCTACTCCTGCCTTTGTACAAGTAAACAAGCTTA Majority
              2460      2470      2480      2490      2500
2451 GCGCACTAATACTCTAGCTACTCCTGCCTTTGTACAAGTAAACAAGCTTA cohl_al2.seq
2451 GCGCACTAATACTCTAGCTACTCCTGCCTTTGTACAAGTAAACAAGCTTA a909_al2.seq

     AGTCCCAATCATTGTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT Majority
              2510      2520      2530      2540      2550
2501 AGTCCCAATCATTGTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT cohl_al2.seq
2501 AGTCCCAATCATTGTCTGATGTGGCAGTTTTATAAACTTTTTCAATCGCT a909_al2.seq

     GTTGGTTCAATAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCCC Majority
              2560      2570      2580      2590      2600
2551 GTTGGTTCAATAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCCC cohl_al2.seq
2551 GTTGGTTCAATAATTTCTCTATTACTGATTTTGTAGTGATAGATTTGCCC a909_al2.seq
```

FIGURE 21C

TGTTGTAGTTGTAAAATAAACATCCGTTCCCATATCTACATTTTTTAAAG Majority
2610 2620 2630 2640 2650
2601 TGTTGTAGTTGTAAAATAAACATCCGTTCCCATATCTACATTTTTTAAAG cohl_ai2.seq
2601 TGTTGTAGTTGTAAAATAAACATCCGTTCCCATATCTACATTTTTTAAAG a909_ai2.seq

CATCAAAATGATAAGGAAAATTATGCGCACAAATCACCATGTTATTAGTT Majority
2660 2670 2680 2690 2700
2651 CATCAAAATGATAAGGAAAATTATGCGCACAAATCACCATGTTATTAGTT cohl_ai2.seq
2651 CATCAAAATGATAAGGAAAATTATGCGCACAAATCACCATGTTATTAGTT a909_ai2.seq

AAATAAGAACCATAATACCTTGTAGGCGTTTTAGACAGTTGTTCAAAACT Majority
2710 2720 2730 2740 2750
2701 AAATAAGAACCATAATACCTTGTAGGCGTCTTAGACAGTTGTTCAAAACT cohl_ai2.seq
2701 AAATAAGAACCATAATACCTTGTAGGCGTTTTAGACAGTTGTTCAAAACT a909_ai2.seq

ATAATTAGCAGCTACCGGTAAATGCAGTTTTAAGTTCGGAATATCCAGAG Majority
2760 2770 2780 2790 2800
2751 ATAATTAGCAGCTACCGGTAAATGCAGTTTTAAGTTCGGAATATCCAGAG cohl_ai2.seq
2751 ATAATTAGCAGCTACCGGTAAATGCAGTTTTAAGTTCGGAATATCCAGAG a909_ai2.seq

TTCCCAAGTAATCTGTTTTATCCAACTTTTTTACAGGTAATTCTCCATTT Majority
2810 2820 2830 2840 2850
2801 TTCCCAAGTAATCTGTTTTATCCAACTTTTTTACAGGTAATTCTCCATTT cohl_ai2.seq
2801 TTCCCAAGTAATCTGTTTTATCCAACTTTTTTACAGGTAATTCTCCATTT a909_ai2.seq

TCTGAACCCTTTACTTGATGCGTAATAGATTTATCAAGCGCCTTGACAAT Majority
2860 2870 2880 2890 2900
2851 TCTGAACCCTTTACTTGATGCGTAATAGATTTATCAAGCACCTTGACAAT cohl_ai2.seq
2851 TCTGAACCCTTTACTTGATGCGTAATAGATTTATCAAGCGCCTTGACAAT a909_ai2.seq

ATGCTGAGAAGTTAAATCAGCTTGATGCGCCTGATTAATATTATACCAAC Majority
2910 2920 2930 2940 2950
2901 ATGCTGAGAAGTTAAATCAGCTTGATGCGCCTGATTAATATTATACCAAC cohl_ai2.seq
2901 ATACTGAGAAGTTAAATCAGCTTGATGCGCCTGATTAATATTATACCAAC a909_ai2.seq

CCCAATAGATTCCAGAACTTACCAGAATGATTCCGAGTATCGCTAAAAAA Majority
2960 2970 2980 2990 3000
2951 CCCAATAGATTCCAGAACTTACCAGAATAATTCCGAGTATAGCTAAAAAA cohl_ai2.seq
2951 CCCAATAGATTCCAGAACTTACCAGAATGATTCCAAGTATCGCTAAAAAA a909_ai2.seq

TTTGCTGAATATCTTCTAATCACGTCTTCTTCTCCATTTTAAGGCTATTA Majority
3010 3020 3030 3040 3050
3001 TTTGCTGAATATCTTCTAATCACGTCTTCTTCTCCATTTTAAAGCTATTA cohl_ai2.seq
3001 TTTGCTGAATATCTTCTAATCACGTCTTCTTCTCCATTTTAAGGCTATTA a909_ai2.seq

TTAAACACAGAAGTCCTGACATAATTAGTATAGGTATTGGCCACCATACT Majority
3060 3070 3080 3090 3100
3051 TTAAACACAGAAGTCCTGACATAATTAGTATAGGTATTGGCCACCATACT cohl_ai2.seq
3051 TTAAACACAGAAGTCCTGACATAATTAGTATAGGTATTGGCCACCATACT a909_ai2.seq

TGTCCAGTAAACGGGAGCTTTCCCTTTGTCTGATGTGTTACTGTAGGAGT Majority
3110 3120 3130 3140 3150
3101 TGTCCAGTAAACGGAAGCTTTCCCTTTGTCTGATGTGTTACTGTAGGAGT cohl_ai2.seq
3101 TGTCCAGTAAACGGGAGCTTTCCCTTTGTCTGATGTGCTACTGTAGGAGT a909_ai2.seq

AATTGTCTCTTCTTTTTTTAGGTTTAGCATTTAAAGGGCTCATTTTCTCAA Majority
3160 3170 3180 3190 3200
3151 AATTGTCTCTTCTTTTTTTAGGTTTAGCATTTAAAGGGCTCATTTTCTCAA cohl_ai2.seq
3151 AATGGTCTCTTCTTTTTTTAGGTTTAGCATTTAAAGGGCTCATTTTCTCAA a909_ai2.seq

ATGCTGTAATATCGTACTTCCCATCCTTAGGTATTGATAGTATAAAGGGA Majority
3210 3220 3230 3240 3250
3201 ATGCTGTAATATCGTACTTCCCATCCTTAGGTATTGATAGTATAAAGGGA cohl_ai2.seq
3201 ATGCTGTAATATCGTACTTCCCATCCTTAGGTATTGATAGTATAAAGGGA a909_ai2.seq

FIGURE 21D

GACATTAGTTCATAACCTTGAGCTGTTTTAGTCTGAATAAATAGATAAAT Majority
3260 3270 3280 3290 3300
3251 GACATTAGTTCATAACCTTGAGCCGTTTTAGTCTGAATAAATAGATAAAT cohl_ai2.seq
3251 GACATTAGTTCATAACCTTGAGCTGTTTTAGTCTGAATAAATAGATAAAT a909_ai2.seq

CCCTTGAGGAAGATTGTTCGCAACAATACCTTCAGCCGGTAAATTATCAA Majority
3310 3320 3330 3340 3350
3301 CCCTTGAGGAAGATTGTTCGCAACAATACCTTCAGCCGGTAAATTATCAA cohl_ai2.seq
3301 CCCTTGAGGAAGATTGTTCGCAACAATACCTTCAGCCGGTAAATTATCAA a909_ai2.seq

ACGTTTGTAAAGGTTGAGTTTTATGAACAGCTTTTGTTAGTAGATTGACG Majority
3360 3370 3380 3390 3400
3351 ACGTTTGTAAAGGTTGAGTTTTATGAACAGCTTTTGTTAGTAGATTGACG cohl_ai2.seq
3351 ACGTTTGTAAAGGTTGAATTTTATGAACAGCTTTTGTTAGTAGATTGACA a909_ai2.seq

TATTTGGCTTGGTTACTATCAAGGTTTACTTGTGTTAGATCATCGTCTTT Majority
3410 3420 3430 3440 3450
3401 TATTTGGCTTGGTTACTATCAAGGTTTACTTGTGTTAGATCATCGTCTTT cohl_ai2.seq
3401 TATTTGGCTTGGTTACTATCAAGGTTTACTTGTGTTAAATCATCGTCTTT a909_ai2.seq

TATTCCAATACCTTGAAATGGGGTAGTTAGAGTAAAAACTTGGTTACCAT Majority
3460 3470 3480 3490 3500
3451 TATTCCAATACCTTGAAATGGGGTAGTTAGAGTAAAAACTTGATTACCAT cohl_ai2.seq
3451 TATTCCAATACCTTGAAATGGGGTAGTTAGAGTAAAAACTTGGTTACCAT a909_ai2.seq

CAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG Majority
3510 3520 3530 3540 3550
3501 CAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG cohl_ai2.seq
3501 CAACATCTTTAGCTTGTGCTACTTGGTAAACAAGTAAATTACCGCCAGCG a909_ai2.seq

ATACCTTGATTATTATACTTATTTTGTATAGTAATAGAACCCGTTTTCAT Majority
3560 3570 3580 3590 3600
3551 ATACCTTGATTATTATACTTATTTTGTATAGTAATAGAACCCGTTTTCAT cohl_ai2.seq
3551 ATACCTTGATTATTATACTTATTTTGTATAGTAATAGAACCCGTTTTCAT a909_ai2.seq

CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA Majority
3610 3620 3630 3640 3650
3601 CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA cohl_ai2.seq
3601 CTGATCATTGGTATCAGCAGACACAAGTTGAGTACTTAGACTAAATAATA a909_ai2.seq

AGAGAAGAGTTATCTTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT Majority
3660 3670 3680 3690 3700
3651 AGAGAAGAGTTATCTTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT cohl_ai2.seq
3651 AGAGAAGAGTTATATTTAGGATCTTTTTATAAATCATTGTTCTCTTCCTT a909_ai2.seq

TCTCATTGCTTGTTTTAAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA Majority
3710 3720 3730 3740 3750
3701 TCTCATTGCTTGTTTTAAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA cohl_ai2.seq
3701 TCTCATTGCTTGTTTTAAAATTTTCTTACGTTGACGTGCTCTCCTAGTTA a909_ai2.seq

CTTCTAAAGAGATTAAAAGTAAAATCAAAGTAAGGAAAATAGCGATAAAT Majority
3760 3770 3780 3790 3800
3751 CTTCTAAAGAGATTAAAAGTAAAATCAAAGTAAGGAAAATAGCGATAAAT cohl_ai2.seq
3751 CTTCTAAAAAGATTAAAAGTAAAATCAAAGTAAGGAAAATAGCGATAAAT a909_ai2.seq

GGTGCGATATAAATAGGCTCTATTTGTATTGCCTCTGCTACTACCAAAGC Majority
3810 3820 3830 3840 3850
3801 GGTGCGATATAAATAGGCTCTATTTGTATTGCCTCTGCTACTACCAAAGC cohl_ai2.seq
3801 GGTGCGATATAAATAGGCTCTATTTGTATTGCCTCTGCTACCACCAAAGC a909_ai2.seq

GTTACCATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG Majority
3860 3870 3880 3890 3900
3851 GTTACCATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG cohl_ai2.seq
3851 GTTACCATTATCGTTTGGTACACGATGTCCTCTCACTAGTAACCGATGGG a909_ai2.seq

FIGURE 21E

```
     TATTAACGCCATATGGTGTACACGTCACCAAAGTTTGGTAGTCTTTACCT Majority
               3910      3920      3930      3940      3950
3901 TATTAACGCCATATGGTGTACACGTCACCAAAGTTTGGTAGTCTTTACCT cohl_al2.seq
3901 TATTAACACCATATGGTGTACACGTCACCAAAGTTTGGTAGTCTTTACCT a909_al2.seq

     TTAACAATTTGTAAATCCCTCAAATCATCCGGTTTAACTGTTCTGATTTG Majority
               3960      3970      3980      3990      4000
3951 TTAACAATTTGTAAATCCCTCAAATCATCCGGTTTAACTGTTCTGATTTG cohl_al2.seq
3951 TTAACAATTTGTAAATCCCTCAAATCATCCGGTTTAACTGTTCTGATTTG a909_al2.seq

     ATCCACTTGATAAGTATATGTTTCATTTAAGATACTGACTGTCCAGTGGT Majority
               4010      4020      4030      4040      4050
4001 ATCCACTTGATAAGTATATGTTTCATTTAAGATACTGACTGTCCAGTGGT cohl_al2.seq
4001 ATCCACTTGATAAGTATATGTTTCATTTAAGATACTGACTGTCCAGTGGT a909_al2.seq

     CTCCAGCTTTTAACTTATCCAAATCAGAAAAAAGCCTTGAAGAGGGTAAA Majority
               4060      4070      4080      4090      4100
4051 CTCCAACTTTTAACTTATCCAAATCAGAAAAAAGCCTTGAAGAGGGTAAA cohl_al2.seq
4051 CTCCAGCTTTTAACTTATCCAAATCAGAAAAAAGCCTTGAAGAGGGTAAA a909_al2.seq

     CCTCTATGTCCTGATAAAATAGAATGAGTTGAGTCTCCTCCAATTGGAAG Majority
               4110      4120      4130      4140      4150
4101 CCTCTATGTCCTGATAAAATAGAATGAGTTGAGTCTCCTCCAATTGGAAG cohl_al2.seq
4101 CCTCTATGTCCTGATAAAATAGAATGAGTTGAGTCTCCTCCAATTGGAAG a909_al2.seq

     ACTACTTCCTTCTAAATGACCAATAGAAGTTTGAAGCACTTTTTCACTTG Majority
               4160      4170      4180      4190      4200
4151 ACTACTTCCTTCTAAATGACCAATAGAAGTTTGAAGCACTTTTTCACTTG cohl_al2.seq
4151 ACTACTTCCTTCTAAATGACCAATAGAAGTTTGAAGCACCTTTTCACTTG a909_al2.seq

     TACCATGATAAAGTGGTAATTTTATGTTTATCTTTGGAATTGAAATATAA Majority
               4210      4220      4230      4240      4250
4201 TACCATGATAAAGTGGTAATTTTATGTTTATCTTTGGAATTGAAATATAA cohl_al2.seq
4201 TACCATGATAAAGTGGTAATTTTATATTTATCTTTGGAATTGAAATATAA a909_al2.seq

     CCCATATTACCCGTTTTATCGATAGCCAGTTGTGAATTATAATCCAAACG Majority
               4260      4270      4280      4290      4300
4251 CCCATATTACCCGTTTTATCGATAGCCAGTTGTGAATTATAATCCAAACG cohl_al2.seq
4251 CCCATATTACCCGTTTTATCGATAGCCAGTTGTGAATTATAATCCAAACG a909_al2.seq

     CTCTTGGTTAGTCATGTGCCACTTCATTCCTGAAGTTTTAAATTGCTTAT Majority
               4310      4320      4330      4340      4350
4301 CTCTTGGCTAGTCATGTGCCACTTCATTCCTGAAGTTTTAAATTGCTTAT cohl_al2.seq
4301 CTCTTGGTCAGTCATGTGCCACTTCATTCCTGAAGTTTTAAATTGCTTAT a909_al2.seq

     TATATTCTTTGGCTCGGTTAATAATTTTTTTATAGTCGTTTTCATCCATA Majority
               4360      4370      4380      4390      4400
4351 TATATTCTTTGGCTCGGTTAATAATTTTTTTATAATCGTTTTCATCCATA cohl_al2.seq
4351 TATATTCTTTGGCTCGGTTAATAATTTTTTTATAGTCGTTTTCATCCATA a909_al2.seq

     TGCGTTACGCGGTCTTGGTAATCCATAATCGCTCGAGATTGGTGAAATGA Majority
               4410      4420      4430      4440      4450
4401 TGCGTTACGCGGTCTTGGTAATCCATAATCGCTCGAGATTGGTGAAATGA cohl_al2.seq
4401 TGCGTTACGCGGTCTTGGTAATCCATAATCGCTCGAGATTGGTGAAATGA a909_al2.seq

     ATTCCAATAATTAGCAAGTGAAGGATAAGCCATTAAGCCTACCCCCACTG Majority
               4460      4470      4480      4490      4500
4451 ATTCCAATAATTAGCAAGTGAAGGATAAGCCATTAAGCCTACCCCCACTG cohl_al2.seq
4451 ATTCCAATAATTAGCAAGTGAAGGATAAGCCATTAAGCCTACCCCCACTG a909_al2.seq

     CAATTATAGTGACAAGCAAAATGGATACTAAATGTTGTCTTATTTTTTTC Majority
               4510      4520      4530      4540      4550
4501 CAATTATAGTGACAAGCAAAATGGATACTAAATGTTGTCTTATTTTTTTC cohl_al2.seq
4501 CAATTATAGTGACAAGCAAAATGGATACTAAATGTTGTCTTATTTTTTTC a909_al2.seq
```

FIGURE 21F

```
     ATATATTTTAAATCTGTACCACTTTGCTAGCCCATCTTATTAAGAACGTA Majority
              4560      4570      4580      4590      4600
4551 ATATATTTTAAATCTGTACCACTTTGCTAGCCCATCTTATTAAGAACGTA coh1_a12.seq
4551 ATATATTTTAAATCTGTACCACTTTGCTAGCCCATCTTATTAAGAACGTA a909_a12.seq

     AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA Majority
              4610      4620      4630      4640      4650
4601 AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA coh1_a12.seq
4601 AACGACGACGAGCAACAAGCACGATACCTGCTCCTATTACTAAAATTGCA a909_a12.seq

     CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC Majority
              4660      4670      4680      4690      4700
4651 CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC coh1_a12.seq
4651 CCTATAATGTAGAAAATTGTTGTACCAATACCACCTGTTGAAGGCAACTC a909_a12.seq

     AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT Majority
              4710      4720      4730      4740      4750
4701 AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT coh1_a12.seq
4701 AGTACCTTTGTTATTTTCAACAGTTGGGTTAACTAAAAGGTTATCTGAAT a909_a12.seq

     TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT Majority
              4760      4770      4780      4790      4800
4751 TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT coh1_a12.seq
4751 TAGTCGTATCAGTGGCTCCATCTCCTAAAATAACCTTCTGAGAGTTATCT a909_a12.seq

     AACAAATTGTAACCTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC Majority
              4810      4820      4830      4840      4850
4801 AACAAATTGTAACCTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC coh1_a12.seq
4801 AACAAATTGTAACCTAAGGGAGCCTTTTTCTCAACTAGATAGTATGTACC a909_a12.seq

     TTCTTTCAAGCCTGTAATGGTAATTATACCATCTGCTCCTGTTGTATATT Majority
              4860      4870      4880      4890      4900
4851 TTCTTTCAAGCCTGTAATGGTAATTATACCATCTGCTCCTGTTGTATATT coh1_a12.seq
4851 TTCTTTCAAGCCTGTAATGGTAATTATACCATCTGCTCCTGTTGTATATT a909_a12.seq

     CTGTTGCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTATCGTTAAAG Majority
              4910      4920      4930      4940      4950
4901 CTGTTGCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTATCGTTAAAG coh1_a12.seq
4901 CTGTTGCATTAGCTTCTGTGCCCCATTCAACGTTATTTGTATCGTTAAAG a909_a12.seq

     TTTAGAAATTGACCCGTAGCATTCTTTAAAACAAATATAGCACCTTGTAA Majority
              4960      4970      4980      4990      5000
4951 TTTAGAAATTGACCCGTAGCATTCTTTAAAACAAATATAGCACCTTGTAA coh1_a12.seq
4951 TTTAGAAATTGACCCGTAGCATTCTTTAAAACAAATATAGCACCTTGTAA a909_a12.seq

     TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC Majority
              5010      5020      5030      5040      5050
5001 TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC coh1_a12.seq
5001 TGAAGCTTTTGTGGAACCATCAATTTTTTTTATAGTAATTTGACCATCCC a909_a12.seq

     TCACTGTTACTTTTTGACCTGGGTCATCATTGCTAGTATTGGGGTTGATG Majority
              5060      5070      5080      5090      5100
5051 TCACTGTTACTTTTTGACCTGGGTCATCATTGCTAGTATTGGGGTTGATG coh1_a12.seq
5051 TCACTGTTACTTTTTGACCTGGGTCATCATTGCTAGTATTGGGGTTGATG a909_a12.seq

     GTCGCAATGTTTGTATTTTCTGGTAAATCAGCTGAACCTGGTTTAGCTCC Majority
              5110      5120      5130      5140      5150
5101 GTCGCAATGTTTGTATTTTCTGGTAAATCAGCTGAACCTGGTTTAGCTCC coh1_a12.seq
5101 GTCGCAATGTTTGTATTTTCTGGTAAATCAGCTGAACCTGGTTTAGCTCC a909_a12.seq

     ACTCTTTAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT Majority
              5160      5170      5180      5190      5200
5151 ACTCTTTAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT coh1_a12.seq
5151 ACTCTTTAATACTCCTGTATAAGTGACTGTGATTGTATTTATTCCCTTAT a909_a12.seq
```

FIGURE 21G

```
     AAAAAAAGTCATCATTAGCTCCATTTTGAGTATTTCCGGTTGGAGTATTG Majority
              5210      5220      5230      5240      5250
5201 AAAAAAAGTCATCATTAGCTCCATTTTGAGTATTTCCGGTTGGAGTATTG cohl_ai2.seq
5201 AAAAAAAGTCATCATTAGCTCCATTTTGAGTATTTCCGGTTGGAGTATTG a909_ai2.seq

     GTAGCTGCCCACGGAATAGTAATCGTGAAATTATTATTTTCCTCTAACAG Majority
              5260      5270      5280      5290      5300
5251 GTAGCTGCCCACGGAATAGTAATCGTGAAATTATTATTTTCCTCTAACAG cohl_ai2.seq
5251 GTAGCTGCCCACGGAATAGTAATCGTGAAATTATTATTTTCCTCTAACAG a909_ai2.seq

     GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAT Majority
              5310      5320      5330      5340      5350
5301 GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAT cohl_ai2.seq
5301 GTTATACTTCCCAGTTGCTTTTTCCGAACCTTGAGTTAGAGTTGTAATAT a909_ai2.seq

     TCCCTGATCCATCAGTAATAGTTACTTCATAAGATCCTTCGTTCAAATCA Majority
              5360      5370      5380      5390      5400
5351 TCCCTGATCCATCAGTAATAGTTACTTCATAAGACCCTTCGTTCAAATCA cohl_ai2.seq
5351 TCCCTGATCCATCAGTAATAGTTACTTCATAAGATCCTTCGTTCAAATCA a909_ai2.seq

     ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT Majority
              5410      5420      5430      5440      5450
5401 ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT cohl_ai2.seq
5401 ACTACAGAAGCAGATGGCATAGTATCCTTTATAACATATTGATACACTTT a909_ai2.seq

     TTCTGTACCATGATAATTGACTGCATTCTTATAAGTAATAGTATATTTGA Majority
              5460      5470      5480      5490      5500
5451 TTCTGTACCATGATAATTGACTGCATTCTTATAAGTAATAGTATATTTGA cohl_ai2.seq
5451 TTCTGTACCATGATAATTGACTGCATTCTTATAAGTAATAGTATATTTGA a909_ai2.seq

     CTGTATCACCAACCGAGTACGTTTTTTGATCTACAGTTTTTCCACCACCA Majority
              5510      5520      5530      5540      5550
5501 CTGTATCACCAACCGAGTACGTTTTTTGATCTACAGTTTTTCCACCACCA cohl_ai2.seq
5501 CTGTATCACCAACCGAGTACGTTTTTTGATCTACAGTTTTTCCACCACCA a909_ai2.seq

     TCTCCCCATGTCGCATCAGTATTCTTTTCATGAATAGTAGCATTTGGAGT Majority
              5560      5570      5580      5590      5600
5551 TCTCCCCATGTCGCATCAGTATTCTTTTCATGAATAGTAGCATTTGGAGT cohl_ai2.seq
5551 TCTCCCCATGTCGCATCAGTATTCTTTTCATGAATAGTAGCATTTGGAGT a909_ai2.seq

     TACAGATGTAACCATAATTACAGCTCCATTATTAACAGTGCTAGAAACAT Majority
              5610      5620      5630      5640      5650
5601 TACAGATGTAACCATAATCACAGCTCCATTATTAACAGTGCTAGAAACAT cohl_ai2.seq
5601 TACAGATGTAACCATAATTACAGCTCCATTATTAACAGTGCTAGAAACAT a909_ai2.seq

     AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT Majority
              5660      5670      5680      5690      5700
5651 AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT cohl_ai2.seq
5651 AATAATATCCATATTGGGAAACATTAATAACCTCAGTACCATCATTATTT a909_ai2.seq

     GACTCAGTAACAGTGGAAACTGGTGTAGTATTAGCTGATATAGATTTAGC Majority
              5710      5720      5730      5740      5750
5701 GACTCAGTAACAGTGGAAACTGGTGTAGTATTAGCTGATATAGATTTAGC cohl_ai2.seq
5701 GACTCAGTAACAGTGGAAACTGGTGTAGTATTAGCTGATATAGATTTAGC a909_ai2.seq

     CCATGTCGCAATCTCATTTGCTGACGCAGTATCTTTTTTAGTTACATATG Majority
              5760      5770      5780      5790      5800
5751 CCATGTCGCAATCTCATTTGCTGACGCAGTATCTTTTTTAGTTACATATG cohl_ai2.seq
5751 CCATGTCGCAATCTCATTTGCTGACGCAGTATCTTTTTTAGTTACATATG a909_ai2.seq

     TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAATTAAAATCAGTTGAA Majority
              5810      5820      5830      5840      5850
5801 TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAATTAAAATCAGTTGAA cohl_ai2.seq
5801 TTCTCCCTCCATTAGTAGTTGTCGTAAAAAGAGAATTAAAATCAGTTGAA a909_ai2.seq
```

FIGURE 21H

```
     GCTTTATACTCAGCTTCTTTACCTTGAGGAATTAAATAAGAAGCTCCATC Majority
              5860      5870      5880      5890      5900
5851 GCTTTATACTCAGCTTCTTTACCTTGAGGAATTAAATAAGAAGCTCCATC cohl_al2.seq
5851 GCTTTATACTCAGCTTCTTTACCTTGAGGAATTAAATAAGAAGCTCCATC a909_al2.seq

     TTTATTCGAATCAGATACATTTGCATTATCTATTTCTGCATCAAAAACTT Majority
              5910      5920      5930      5940      5950
5901 TTTATTCGAATCAGATACATTTGCATTATCTATTTCTGCATCAAAAACTT cohl_al2.seq
5901 TTTATTCGAATCAGATACATTTGCATTATCTATTTCTGCATCAAAAACTT a909_al2.seq

     TGTATGCTTTATAGGTTGCGCCTTTTTGAGTATCTTGAACTGTAATTGTC Majority
              5960      5970      5980      5990      6000
5951 TATATGCTTTATAGGTTGCGCCTTTTTGAGTATCTTGAACTGTAATTGTC cohl_al2.seq
5951 TGTATGCTTTATAGGTTGCGCCTTTTTGAGTATCTTGAACTGTAATTGTC a909_al2.seq

     CCTGTCTCAGCGGCAAAAGCTATCGGCGTAACTGGTGATACAGCCATACC Majority
              6010      6020      6030      6040      6050
6001 CCTGTCTCAGCGGCAAAAGCTATCGGCGTAACTGGTGATACAGCCATACC cohl_al2.seq
6001 CCTGTCTCAGCGGCAAAAGCTATCGGCGTAACTGGTGATACAGCCATACC a909_al2.seq

     AAATGCTAAACTCGCCACTAACAGCGATTGAATCATTTTCTTTTTCATTG Majority
              6060      6070      6080      6090      6100
6051 AAATGCTAAACTCGCCACTAACAGCGATTGAATCATTTTCTTTTTCATTG cohl_al2.seq
6051 AAATGCTAAACTCGCCACTAACAGCGATTGAATCATTTTCTTTTTCATTG a909_al2.seq

     AAATCTTTCTCCTAAAATCATATTGATGAATGATTAATTCATATTTTTTT Majority
              6110      6120      6130      6140      6150
6101 AAATCTTTCTCCTAAAATCATATTGATGAATGATTAATTCATATTTTTTT cohl_al2.seq
6101 AAATCTTTCTCCTAAAATCATATTGATGAATGATTAATTCATATTTTTTT a909_al2.seq

     TCGATAGTATAATATTAATCCTGATGGTAGAGCTAAAGCTAAACCAACTA Majority
              6160      6170      6180      6190      6200
6151 TCGATAGTATAATATTAATCCTGATGGTAGAGCTAAAGCTAAACCAACTA cohl_al2.seq
6151 TCGATAGTATAATATTAATCCTGATGGTAGAGCTAAAGCTAAACCAACTA a909_al2.seq

     GGATATAAATGTGTGTTCCAATACCTCCAGTACTAGGCAATTCTGTTCCT Majority
              6210      6220      6230      6240      6250
6201 GGATATAAATGTGTGTTCCAATACCTCCAGTACTAGGCAATTCTGTTCCT cohl_al2.seq
6201 GGATATAAATGTGTGTTCCAATACCTCCAGTACTAGGCAATTCTGTTCCT a909_al2.seq

     TTACTGTTAGTAATTTTAAAAGTATATACTGTACTTCCATCTACTAAATT Majority
              6260      6270      6280      6290      6300
6251 TTACTGTTAGTAATTTTAAAAGTATATACTGTACTTCCATCCACTAAATT cohl_al2.seq
6251 TTACTGTTAGTAATTTTAAAAGTATATACTGTACTTCCATCTACTAAATT a909_al2.seq

     CTCTTTTATTGGTGTCGCATTATTACCATTTTGTTCAAAGGTAACTCCCG Majority
              6310      6320      6330      6340      6350
6301 CTCTTTTATTGGTGTCGCATTATTACCATTTTGTTCAAAGGTAACTCCCG cohl_al2.seq
6301 CTCTTTTATTGGTGTCGCATTATTACCATTTTGTTCAAAGGTAACTCCCG a909_al2.seq

     TAGAAATCACTAATACTGATATATCATTTTTAGGTAGTAGGTACCCTGGA Majority
              6360      6370      6380      6390      6400
6351 TAGAAATCACTAATACTGATATATCATTTTTAGGTAGTAGGTACCCTGGA cohl_al2.seq
6351 TAGAAATCACTAATACTGACATATCATTTTTAGGTAGTAGGTACCCTGGA a909_al2.seq

     GGGGCCTTTGTCTCTGTTAGGTAGTATTTTCCTACTGGCAAACTGAGGTA Majority
              6410      6420      6430      6440      6450
6401 GGGGCCTTTGTCTCTGTTAGGTAGTATTTTCCTACTGGCAAACTGAGGTA cohl_al2.seq
6401 GGGGCCTTTGTCTCTGTTAGGTAGTATTTTCCTACTGGCAAACTGAGGTA a909_al2.seq

     GTTATTAGCATCCACTAATAACAAGCCTTTATCGTTTGTCACCAGCCCTG Majority
              6460      6470      6480      6490      6500
6451 GTTATTAGCATCCACTAATAACAAGCCTTTATCGTTTGTCACCAGCCCTG cohl_al2.seq
6451 GTTATTAGCATCCACTAATAACAAGCCTTTATCGTTTGTCACCAGCCCTG a909_al2.seq
```

FIGURE 21I

```
     AATACATAGGATGTGAAGCTTTATTCCCATTAGCATCTGATTCATAAATA Majority
              6510      6520      6530      6540      6550
6501 AATACATAGGATGTGAAGCTTTATTCCCATTAGCATCTGATTCATAAATA coh1_ai2.seq
6501 AATACATAGGATGTGAAGCTTTATTCCCATTAGCATCTGATTCATAAATA a909_ai2.seq

     TCAAAAACTGCACCTGCTAAAAAATTATTATCATTTTCGACATTAACTTT Majority
              6560      6570      6580      6590      6600
6551 TCAAAAACTGCACCTGCTAAAAAATTATTATCATTTTCGACATTAACTTT coh1_ai2.seq
6551 TCAAAAACTGCACCTGCTAAAAAATTATTATCATTTTCGACATTAACTTT a909_ai2.seq

     CTGTAGTCGTACTTTTTGCTTGATACGTGTATTGGTAAAGCTAATATCTA Majority
              6610      6620      6630      6640      6650
6601 CTGTAGTCGTACTTTTTGCTTGATACGTGTATTGGTAAAGCTAATATCTA coh1_ai2.seq
6601 CTGTAGTCGTACTTTTTGCTTGATACGTGTATTGGTAAAGCTAATATCTA a909_ai2.seq

     CGTCTCCTGAAACTGTCAGGGATTGTAAGCCGGTAGCATCATAAGTTTTA Majority
              6660      6670      6680      6690      6700
6651 CGTCTCCTGAAACTGTCAGGGATTGTAAGCCGGTAGCATCATAAGTTTTA coh1_ai2.seq
6651 CGTCTCCTGAAACTGTAAGGGATTGTAAGCCGGTAGCATCATAAGTTTTA a909_ai2.seq

     TCAGCTTCACCAGTTGCTAGATTTTTTTCTGTAATTGACTCAGATACTTT Majority
              6710      6720      6730      6740      6750
6701 TCAGCTTCACCAGTTGCTAGATTTTTTTCTGTAATTGACTCAGATACTTT coh1_ai2.seq
6701 TCAGCTTCACCAGTTGCTAGATTTTTTTCTGTAATTGACTCAGATACTTT a909_ai2.seq

     AAATTCATCGTAGGCTTGTTCATCTATTGATATAGAAGTTCCATAAGGTA Majority
              6760      6770      6780      6790      6800
6751 AAATTCATCGTAGGCTTGTTCATCTATTGATATAGAAGTTCCATAAGGTA coh1_ai2.seq
6751 AAATTCATCGTAGGCTTGTTCATCTATTGATATAGAAGTTCCATAAGGTA a909_ai2.seq

     CTTTAAATTCCTTAGTCTGACCATCTCTCAGCGGAAAATTCTCTTGTTGC Majority
              6810      6820      6830      6840      6850
6801 CTTTAAATTCCTTAGTCTGACCATCTCTCAGCGGAAAATTCTCTTGTTGC coh1_ai2.seq
6801 CTTTAAATTCCTTAGTCTGACCATCTCTCAGCGGAAAATTCTCTTGTTGC a909_ai2.seq

     AACGTTTCACTTGGATTAAACAAGAAGTCTTTCGTCTTATCTTCATCTAG Majority
              6860      6870      6880      6890      6900
6851 AACGTTTCACTTGGATTAAACAAGAAGTCTTTCGTCTTATCTTCATCTAG coh1_ai2.seq
6851 AACGTTTCACTTGGATTAAACAAGAAGTCTTTCGTCTTATCTTCATCTAG a909_ai2.seq

     TCCAACGACAGTTTTACTTACTCTGACGGTGTATTCTTTAGGTTGCCAAA Majority
              6910      6920      6930      6940      6950
6901 TCCAACGACAGTTTTACTTACTCTGACGGTGTATTCTTTAGGTTGCCAAA coh1_ai2.seq
6901 TCCAACGACAGTTTTACTTACTCTGACGGTGTATTCTTTAGGTTGCCAAA a909_ai2.seq

     CAGCATATAAGGTATTTGTTGCATCAGGGTTGTTATCAATACCTATTGAT Majority
              6960      6970      6980      6990      7000
6951 CAGCATATAAGGTATTTGTTGCATCAGGGTTGTTATCAATACCTATTGAT coh1_ai2.seq
6951 CAGCATATAAGGTATTTGTTGCATCAGGGTTGTTATCAATACCTATTGAT a909_ai2.seq

     TGACCTGCTGTAAATTCCACACGTCCTGTATCAGCTAAATCCTTATCATG Majority
              7010      7020      7030      7040      7050
7001 TGACCTGCTGTAAATTCCACACGTCCTGTATCAGCTAAATCCTTATCATG coh1_ai2.seq
7001 TGACCTGCTGTAAATTCCACACGTCCTGTATCAGCTAAATCCTTATCATG a909_ai2.seq

     ATGCCAACCAATAAGGTTGTAACCTGTCCTTGTAAAGTATTGGTTTTCAG Majority
              7060      7070      7080      7090      7100
7051 ATGCCAACCAATAAGGTTGTAACCTGTCCTTGTAAAGTATTGGTTTTCAG coh1_ai2.seq
7051 ATGCCAACCAATAAGGTTGTAACCTGTCCTTGTAAAGTATTGGTTTTCAG a909_ai2.seq

     GAATTGTAGTTGTGCTATTCAACTCCATACGCGGTGTCTCTACTTGTGTT Majority
              7110      7120      7130      7140      7150
7101 GAATTGTAGTTGTGCTATTCAACTCCATACGCGGTGTCTCCACTTGTGTT coh1_ai2.seq
7101 GAATTGTAGTTGTGCTATTCAACTCCATACGCGGTGTCTCTACTTGTGTT a909_ai2.seq
```

FIGURE 21J

```
     ACCACATTACCATTTTCTACTCTAGTACCACCGTTACCATTGTATTTGAT Majority
              7160      7170      7180      7190      7200
7151 ACCACATTACCATTTTCTACTCTAGTACCACCGTTACCATTGTATTTGAT cohl_al2.seq
7151 ACCACATTACCATTTTCTACTCTAGTACCACCGTTACCATTATATTTGAT a909_al2.seq

     TGAGGTATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA Majority
              7210      7220      7230      7240      7250
7201 GGAGGTATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA cohl_al2.seq
7201 TGAGGTATCTTCTAATTTGATATCTCCTACTGGAATAATGACAGGTTTTA a909_al2.seq

     TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA Majority
              7260      7270      7280      7290      7300
7251 TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA cohl_al2.seq
7251 TGGTGATATTTTTATTAGCATCTGCTAAATGGGCGTCAATATCAATGGAA a909_al2.seq

     TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA Majority
              7310      7320      7330      7340      7350
7301 TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA cohl_al2.seq
7301 TCATATGGGTTATAAATTTTACCATTGTACCACCAGCCACGGAAACGATA a909_al2.seq

     GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT Majority
              7360      7370      7380      7390      7400
7351 GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT cohl_al2.seq
7351 GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT a909_al2.seq

     CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTTTGTAATTGCTGACCA Majority
              7410      7420      7430      7440      7450
7401 CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTTTGTAATTGCTGACCA cohl_al2.seq
7401 CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTTTGTAATTGCTGACCA a909_al2.seq

     GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA Majority
              7460      7470      7480      7490      7500
7451 GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA cohl_al2.seq
7451 GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATTGCTGTA a909_al2.seq

     TATAATATGATAATCTCCAGCCTTTCGCCAAATAGCTCTTAAATTGATAT Majority
              7510      7520      7530      7540      7550
7501 TATAATATGATAATCTCCAGCCTTTCGCCAAATAGCTCTTAAATTGATAT cohl_al2.seq
7501 TATAATATGATAATCTCCAGCCTTTCGCCAAATAGCTCTTAAATTGATAT a909_al2.seq

     CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA Majority
              7560      7570      7580      7590      7600
7551 CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA cohl_al2.seq
7551 CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGGA a909_al2.seq

     TTAACATAATACCAACCGACTAATTTGTAAGCGTCTTTTACGTACTTGTA Majority
              7610      7620      7630      7640      7650
7601 TTAACATAATACCAACCGACTAATTTGTAAGCGTCTTTTACGTACTTGTA cohl_al2.seq
7601 TTAACATAATACCAACCGACTAATTTGTAAGCGTCTTTTACGTACTTGTA a909_al2.seq

     CTTAGTTGTTGTATCAACATTTGAGAGACTAGTATCTGTCGTATAATAGG Majority
              7660      7670      7680      7690      7700
7651 CTTAGTTGTTGTATCAACATTTGAGAGACTAGTATCTGTCGTATAATAGG cohl_al2.seq
7651 CTTAGTTGTTGTATCAACATTTGAGAGACTAGTATCTGTCGTATAATAGG a909_al2.seq

     CATCTTTAGTTGAGTCGGGATCTTTATCTCGTGAATCATACTTATAATAA Majority
              7710      7720      7730      7740      7750
7701 CATCTTTAGTTGAGTCGGGATCTTTATCTCGTGAATCATACTTATAATAA cohl_al2.seq
7701 CATCTTTAGTTGAGTCGGGATCTTTATCTCGTGAATCATACTTATAATAA a909_al2.seq

     TATGTACCTGAAGGATCTTGGATATAATCCCTTGTAATATCTGTATAATC Majority
              7760      7770      7780      7790      7800
7751 TATGTACCTGAAGGATCTTGGATATAATCCCTTGTAATATCTGTATAATC cohl_al2.seq
7751 TATGTACCTGAAGGATCTTGGATATAATCCCTTGTAATATCTGTATAATC a909_al2.seq
```

FIGURE 21K

```
      CGGAATACGATCACCATAATGCAAGTCTAAATAGGTATCATCTGTTTTTG Majority
               7810      7820      7830      7840      7850
7801  CGGAATACGATCACCATAATGCAAATCTAAATAGGTATCATCTGTTTTTG cohl_ai2.seq
7801  CGGAATACGATCACCATAATGCAAGTCTAAATAGGTATCATCTGTTTTTG a909_ai2.seq

      ATAATTGGCCTCCGTTTGGATCAATATTGACACGATATGTTACCTTTTGC Majority
               7860      7870      7880      7890      7900
7851  ATAATCGGCCTCCGTTTGGATCAATATTGACACGATATGTTACCTTTTGC cohl_ai2.seq
7851  ATAATTGGCCTCCGTTTGGATCAATATTGACACGATATGTTACCTTTTGC a909_ai2.seq

      CAACCTGCATAGACTTTAACATCATGAGGAGGCATAGTCGTGTTAAAGTC Majority
               7910      7920      7930      7940      7950
7901  CAACCTGCATAGACTTTAACATCATGAGGAGGCATAGTCGTGTTAAAGTC cohl_ai2.seq
7901  CAACCTGCATAGACTTTAACATCATGAGGAGGCATAGTCGTGTTAAAGTC a909_ai2.seq

      AAATACTTGTGTTTGTGCTTGGTCTTTATACCATTTACCATCCCAAACAT Majority
               7960      7970      7980      7990      8000
7951  AAATACTTGTGTTTGTGCTTGGTCTTTATACCATTTACCATCCCAAACAT cohl_ai2.seq
7951  AAATACTTGTGTTTGTGCTTGGTCTTTATACCATTTACCATCCCAAACAT a909_ai2.seq

      ACCCTGGTCGACTAGGTTTAGGTTGAACCGTTGTCGTATCGGGGGCATAA Majority
               8010      8020      8030      8040      8050
8001  ACCCTGGTCGACTAGGTTTAGGTTGAACAGTTGTCGTATCGGGGGCATAA cohl_ai2.seq
8001  ACCCTGGTCGACTAGGTTTAGGTTGAACCGTTGTCGTATCGGGGGCATAA a909_ai2.seq

      GAGGACAAATTTTGCTCATATAGAACATCCTTTACTGGAAAATTAGGAAG Majority
               8060      8070      8080      8090      8100
8051  GAGGACAAATTTTGCTCATATAGAACATCCTTTACTGGAAAATTAGGAAG cohl_ai2.seq
8051  GAGGACAAATTTTGCTCATATAGAACATCCTTTACTGGAAAATTAGGAAG a909_ai2.seq

      CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT Majority
               8110      8120      8130      8140      8150
8101  CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT cohl_ai2.seq
8101  CTCTGTATTATCAAGCGGATCTAAATATTTAATCTTGTATGAATTACGTT a909_ai2.seq

      CATACCATACCACTAAGTTCAAATAATCTTTGTGGTAGTCTCCATATTTA Majority
               8160      8170      8180      8190      8200
8151  CATACCATACCACTAAGTTCAAATAATCTTTGTGGTAGGCACCATATCTA cohl_ai2.seq
8151  CATACCATACCACTAAGTCCAAATAATCTTTGTGATAATATCCATACTGA a909_ai2.seq

      TCGTAGTATTCATCTGCGATTGGCACTTTTGTTTTTGCACTCGTTTGTCT Majority
               8210      8220      8230      8240      8250
8201  TCATAGTATTCATCTGAAATAGGAACTTTTACTCCTGCACTCGTTTGACT cohl_ai2.seq
8201  TCGTAATATTCATCAGCGAGTGGCACAGTTGTATTAGCACTAGTTTGTAG a909_ai2.seq

      TGGGTTCTGATCAAATAGGTAATTATCTGGATATAAGCTTTGATAGTATT Majority
               8260      8270      8280      8290      8300
8251  TGCGTACTGATCAAAAAGGTAATTATCAGGATATAAACGTTGATAATATT cohl_ai2.seq
8251  TGGCTTAGCATCAAATAAGTAATTATCTGGATACAAGCTTTGATAGTATT a909_ai2.seq

      TAACATTAAATCCTAGGTATTTTTCTGTAAAGGTAAATTCGTCTGGTCCA Majority
               8310      8320      8330      8340      8350
8301  TAACATTAAACCCTAGATATTTTTCTGTAAAGGTAAATTCATCTGCCCCA cohl_ai2.seq
8301  TAACATTAAATCCTAAGTATTTTTCTGTAAAGGTAAACTCGTCTGCTCCA a909_ai2.seq

      GCACCTCCCCCTGTGTCTGCTAAAGAGTATTTGCCATCTAGTCCTTGTTT Majority
               8360      8370      8380      8390      8400
8351  GCACCTCCACCGGTATCTGCTAAAGAATAAGTGCCATCCAAACCTTGTTT cohl_ai2.seq
8351  GCACCTCCCCCTACGTCTGATAAAAAGTATTTGCCATCTAGTCCTTGTTT a909_ai2.seq

      GTAGAACGGATAATTTTGAATTCTCTTCCCTTTTGGATAGAGTTTTATTT Majority
               8410      8420      8430      8440      8450
8401  GAAGAACCGATAAGTTTGAACTATATTCCCTTTTGGAAAGAGTTTGATAA cohl_ai2.seq
8401  ATAGAAAGCATAATTTTGAATTCGCTGACCAACAGGATACAATTTTATTT a909_ai2.seq
```

FIGURE 21L

```
     CATCTGGATTTGCTCTATGTACTCCATTGGGTAGTATGAACTCACCCAAA Majority
            8460      8470      8480      8490      8500
8451 CATCTGAATCAACGGTAT-----CATTAGGTAAAATGAACTCACCCAAA cohl_al2.seq
8451 CATAAGGATTTGCTCCAAGTACTCCAGTGGAAGTATAAACTCACCCAAA a909_al2.seq

     TAACTCATTCCTTATGTTCCAGTTTGGTTATTTCCAACATTGGTTAGGTA Majority
            8510      8520      8530      8540      8550
8495 TAACTCATCCCATAGGTTCCAACTTGGTTATTTCCAACATTGGCTAAAATA cohl_al2.seq
8501 TAACTCATTCCTGATACTCCAGTTGAA-----------------GTTAGGTA a909_al2.seq

     ACGCCATGCACCTGTCTTCCATTGATAGCCATTCGCGGCTAAGGTTGTAC Majority
            8560      8570      8580      8590      8600
8545 ACCCCATGCACCGGTCTTCCATTGATAACCATTAGCGGCCAAGGTTGTAC cohl_al2.seq
8536 ACGCCAAGCACCTGTCTTCCAATGATAGCCATTCGCAGCTAAAGTCGTAC a909_al2.seq

     CGTATAGTCCTGTGTAGGTTTCGGCATCTGATGCTCTAGTTCTAGGAATC Majority
            8610      8620      8630      8640      8650
8595 CGTAAAGTCCTGTGTAAGTATCAGCATCAGA-GCTCCAATTATACGAATA cohl_al2.seq
8586 CATATAGTCCAGTATAGGTTTCGGCATATGATGCACTAGTCCAAGGAAAC a909_al2.seq

     GTAGTATTTTGCTAATGAATCTCCGAGTAGCCCTTTTTTGCAAATTTTAT Majority
            8660      8670      8680      8690      8700
8644 GTAATATTCTGCCAAGGAATATCCG------CCATATTTTGCAAATTTTAT cohl_al2.seq
8636 GAAGAACTTTGATAATAAACCTCAGAGTAGCCCGTCTTCGCAAATCTCAT a909_al2.seq

     TGTGATGAGTTTTCTATCATAATAAACATTAACGACACTTGAACCATCGT Majority
            8710      8720      8730      8740      8750
8689 TGTGATGAGTTTTCTATCATAATAAACATTAACAACACTTGAACCATCGT cohl_al2.seq
8686 AGTCACAAGTTCTCTATCATAATAAACATTAACGACACTTGAACCATCGT a909_al2.seq

     CTTTTATCATGACAGAAGTTTCTGTCCTCGTATTATTAACTTTAAAGCCA Majority
            8760      8770      8780      8790      8800
8739 CTTTTATCATGACAGAAGTTTCTGTCCTCGTATTATTAACTTTAAAGCCA cohl_al2.seq
8736 CTTTTATCATGACAGAAGTTTCTGTCCTCGTATTATTAACTTTAAAGCCA a909_al2.seq

     GTCGGTAGTTTTTCATTAATATCTTGTTGTGTTAGCGTCTGATTAGATAA Majority
            8810      8820      8830      8840      8850
8789 GTCGGTAGTTTATCATTAATATCTTGTTGTGTTAGCGTCTGATTAGATAA cohl_al2.seq
8786 GTCGGCAATTTTGCATTAATATCTTGTTGTGTTAGCGTCTGATTAGATAA a909_al2.seq

     AGATAGGCCTGATCGTGTTACTTGCCCTGCGTACTCATATGTCTTTTGCG Majority
            8860      8870      8880      8890      8900
8839 AGATAGGCCTGATCGTGTTACTTGCCCTGCGTACTCATATGTCTTTTGCG cohl_al2.seq
8836 AGATAGGCCTGATCGTGTTACTTGCCCTGCGTACTCATATGTCTTTTGCG a909_al2.seq

     CATCAGTAGCATTTTTATTATCCGTTGCTGATTGTTGCCAGTAGTTTATC Majority
            8910      8920      8930      8940      8950
8889 CATCAGTAGCATTTTTATTATCCGTTGCTGATTGTTGCCAGTAGTTTATC cohl_al2.seq
8886 CATCAGTAGCATTTTTATTATCCGTTGCTGATTGTTGCCAGTAGTTTATC a909_al2.seq

     GTGTAGGTTGTTTGTGCGGGCGACCAATGTGCATATAGCGTCGTATCCTT Majority
            8960      8970      8980      8990      9000
8939 GTGTAGGTTGTTTGTGCGGGCGACCAATGTGCATATAGCGTCGTATCCTT cohl_al2.seq
8936 GTGTAGGTTGTTTGTGCGGGCGACCAATGTGCATATAGCGTCGTATCCTT a909_al2.seq

     GGTCAAGACTTGATTAAAATCAAAGGCTGTCCCACCACTAGCAGCTGTGT Majority
            9010      9020      9030      9040      9050
8989 GGTCAAGACTTGATTAAAATCAAAGGCTGCCCCACCACTAGCAGCTGTGT cohl_al2.seq
8986 GGTCAAGACTTGATTAAAATCAAAGGCTGTCCCACCACTAGCAGCTGTGT a909_al2.seq

     ACCACCCTGCAAAAGTATAACCTGGCCTTGTTGGATCATTAGGCTTAATT Majority
            9060      9070      9080      9090      9100
9039 ACCACCCTGCAAAAGTATAACCTGGCCTCGTTGGATCATTAGGCTTAATT cohl_al2.seq
9036 ACCACCCTGCAAAAGTATAACCTGGCCTTGTTGGATCATTAGGCTTAATT a909_al2.seq
```

FIGURE 21M

```
     GTCGAAGCAGGTTGGTCTGTTAACACACGACGAGGTGCAATATAGGTAAC Majority
               9110      9120      9130      9140      9150
9089 GTCGAAGCAGGTTGGCCTGTTAACACACGACGAGGTGCAATATAGGTAAC cohl_ai2.seq
9086 GTCGAAGCAGGTTGGTCTGTTAACACACGACGAGGTGCAATATAGGTAAC a909_ai2.seq

     TCCTGTTGATAAGTCGCCTGTGTTGAATTCAACACCGTACGATTCTTTAA Majority
               9160      9170      9180      9190      9200
9139 TCCTGTTGATAAGTCGCCTGTGTTGAATTCAACACCGTACGATTCTTTAA cohl_ai2.seq
9136 TCCTGTTGATAAGTCGCCTGTGTTGAATTCAACACCGTACGATTCTTTAA a909_ai2.seq

     AGACAGGATAAAGACTTATGTCTGAAGATACAGGTAATGCTTGAATTTCT Majority
               9210      9220      9230      9240      9250
9189 AGACAGGATAAAGACTTATGTCTGAAGATACAGGTAATGCTTGAATTTCT cohl_ai2.seq
9186 AGACAGGATAAAGACTTATGTCTGAAGATACAGGTAATGCTTGAATTTCT a909_ai2.seq

     GATTCAGAAAGTGGAGCTCCATTTTGAGTTTTAGACCAACCTACAAATAA Majority
               9260      9270      9280      9290      9300
9239 GATTCAGAAAGTGGAGCTCCATTTTGAGTTTTAGACCAACCTACAAATAA cohl_ai2.seq
9236 GATTCAGAAAGTGGAGCTCCATTTTGAGTTTTAGACCAACCTACAAATAA a909_ai2.seq

     TAATGTTGAAGTGGGTGGTGAAACTTTAAATGAACTTAGATCCGCGGTTC Majority
               9310      9320      9330      9340      9350
9289 CAATGTTGAAGTGGGTGGTGAAACTTTAAATGAACTTAAATCCGCGGTTC cohl_ai2.seq
9286 TAATGTTGAAGTGGGTGGTGAAACTTTAAATGAACTTAGATCCGCGGTTC a909_ai2.seq

     CGTTTCCTGAACTATCCAATGGTACTTGATAACGCTCCAAAATACTTTTA Majority
               9360      9370      9380      9390      9400
9339 CGTTTCCTGAACTATCCAATGGTACTTGATAACGCTCCAAAATACTTTTA cohl_ai2.seq
9336 CGTTTCCTGAACTATCCAATGGTACTTGATAACGCTCCAAAATACTTTTA a909_ai2.seq

     CCAGCAGAATCATCATATAAAGTTACTGTCGCCACTTTACTCATAAATGG Majority
               9410      9420      9430      9440      9450
9389 CCAGCAGAATCATCATATAAAGTTACTGTCGCCACTTTACTCATAAATGG cohl_ai2.seq
9386 CCAGCAGAATCATCATATAAAGTTACTGTCGCCACTTTACTCATAAATGG a909_ai2.seq

     ACGAACATAAATTTCTTTTGTCTCAGTTACAGTTATTGGCTCACCAAATT Majority
               9460      9470      9480      9490      9500
9439 ACGAACACAAATTTCTTTTGTCTCAGTTACAGTTATTGGCTCACCAAATT cohl_ai2.seq
9436 ACGAACATAAATTTCTTTTGTCTCAGTTACAGTTATTGGCTCACCAAATT a909_ai2.seq

     TAACAGGGTCACCATACTTTCCAGTAGTAGGATCATAGGTATACCAACCA Majority
               9510      9520      9530      9540      9550
9489 TAACAGGGTCACCATACTTTCCAGTAGTAGGATCATAGGTATACCAACCA cohl_ai2.seq
9486 TAACAGGGTCACCATACTTTCCAGTAGTAGGATCATAGGTATACCAACCA a909_ai2.seq

     TTAAAATGCTCTCCTGCTTTAATCGTCGGAATCCCAACTTCTCCTAGAGA Majority
               9560      9570      9580      9590      9600
9539 TTAAAATGCTCTCCTGCTTTAATCGTCGGAATCCCAACTTCTCCTAGAGA cohl_ai2.seq
9536 TTAAAATGCTCTCCTGCTTTAATCGTCGGAATCCCAACTTCTCCTAGAGA a909_ai2.seq

     TTCTCCATCTTTTATAATTTGATGATGAACTTGCATACCTGAAGCTGTCA Majority
               9610      9620      9630      9640      9650
9589 TTCTCCATCTTTTATAATTTGATGATGAACTTGCATACCTGAAGCTGTCA cohl_ai2.seq
9586 TTCTCCATCTTTTATAATTTGATGATGAACTTGCATACCTGAAGCTGTCA a909_ai2.seq

     GGAAATTATAATCAGTTCCGTCATTATTTTGAAAATGGTAAGTTAACCTA Majority
               9660      9670      9680      9690      9700
9639 GGAAATCATAATCAGTTCCGTCATTATTTTGAAAATGGTAAGTTAACCTA cohl_ai2.seq
9636 GGAAATTATAATCAGTTCCGTCATTATTTTGAAAATGGTAAGTTAACCTA a909_ai2.seq

     GGAACTTCTGTATTATCCTCTTGAACAATTGCATAAATGGAGAATGAATC Majority
               9710      9720      9730      9740      9750
9689 GGAACTTCTGTATTATCCTCTTGAACAATTGCATAAATGGAGAATGAATC cohl_ai2.seq
9686 GGAACTTCTGTATTATCCTCTTGAACAATTGCATAAATGGAGAATGAATC a909_ai2.seq
```

FIGURE 21N

```
      TGTTTTAAAGCAACATCACTGCTAGTGTTCTTAGTTTCTGCAGTATCTT Majority
              9760      9770      9780      9790      9800
9739  TGTTTTAAAGCAACATCACTGCTAGTGTTCTTAGTTTCTGCAGTATCTT coh1_ai2.seq
9736  TGTTTTAAAGCAACATCACTGCTAGTGTTCTTAGTTTCTGCAGTATCTT a909_ai2.seq

      TAGATTTTAATACTTCTGTTTGACCATCATCTTTAAAGTGAACAACTTTA Majority
              9810      9820      9830      9840      9850
9789  TAGATTTTAATACTTCTGTTTGACCATCATCTTTAAAGTGAACAACTTTA coh1_ai2.seq
9786  TAGATTTTAATACTTCTGTTTGACCATCATCTTTAAAGTGAACAACTTTA a909_ai2.seq

      AGGTTTTCATCTGAAGCTTCTAATGGCTTATCATAGTTGACCTCTACTTT Majority
              9860      9870      9880      9890      9900
9839  AGGTTTTCATCTGAAGCTTCTAATGGCTTATCATAGTTAACCTCTACTTT coh1_ai2.seq
9836  AGGTTTTCATCTGAAGCTTCTAATGGCTTATCATAGTTGACCTCTACTTT a909_ai2.seq

      TACTGGGGCTTGGGGTTCTGCTTCTTTACCATTTGACTCAATAGTAATGT Majority
              9910      9920      9930      9940      9950
9889  TACTGGAGCTTGGGGTTCTGCTTCTTTACCATTTGACTCAATAGTAATGT coh1_ai2.seq
9886  TACTGGGGCTTGGGGTTCTGCTTCTTTACCATTTGACTCAATAGTAATGT a909_ai2.seq

      CATAGAGTTTGAAGTTTTTGATTTCACTATCTTGTTTAGCAACTTCTGTC Majority
              9960      9970      9980      9990     10000
9939  CATAGAGTTTGAAATTTTTGATTTCACTATCTTGTTTAGCAACCTCTGTC coh1_ai2.seq
9936  CATAGAGTTTGAAGTTTTTGATTTCACTATCTTGTTTAGCAACTTCTGTC a909_ai2.seq

      AATGCTTTTTTCTTATAGTCTTTAAAAGTAGCTGAATTGTCTTTTAATTC Majority
             10010     10020     10030     10040     10050
9989  AATGCTTTTTTCTTATAATCTTTAAAAGTAGCTGAATTGTCTTTTAATTC coh1_ai2.seq
9986  AATGCTTTTTTCTTATAGTCTTTAAAAGTAGCTGAATTGTCTTTTAATTC a909_ai2.seq

      CGTCACCTTTAAATCAGCATTTTTAGGAATCTTAGCTTCTTTGGTCAAAG Majority
             10060     10070     10080     10090     10100
10039 CGTCACCTTTAAATCAGCATTTTTAGGAATCTTAGCTTCTTTGGTCAAAG coh1_ai2.seq
10036 CGTCACCTTTAAATCAGCATTTTTAGGAATCTTAGCTTCTTTGGTCAAAG a909_ai2.seq

      TCACTGTTACAGTATAGTCTGCACCTCTAAACATCAATGGTTCTTCACGG Majority
             10110     10120     10130     10140     10150
10089 TCACTGTTACAGTATAGTCTGCACCTCTAAACATCAATGGTTCTTCACGA coh1_ai2.seq
10086 TCACTGTTACAGTATAGTCTGCACCTCTAAACATCAATGGTTCTTCACGG a909_ai2.seq

      TAAGCAGCTTCCTCAGAAGATGATGTTTCTGTTACACTAGAAGCAGGAGT Majority
             10160     10170     10180     10190     10200
10139 TAAGCAGCTTCCTCAGAAGATGATGTTTCTGTTACACTAGAAGCAGGAGT coh1_ai2.seq
10136 TAAGCAGCTTCCTCAGAAGATGATGTTTCTGTTACACTAGAAGCAGGAGT a909_ai2.seq

      CTGTGGCTTGCTCTGCTCAACACTTGATTGAGAACTAGATGTTGATGAAG Majority
             10210     10220     10230     10240     10250
10189 CTCGGGCTTGCTCTGCTCAACACTTGATTGAGAACTAGATGTTGATGAAG coh1_ai2.seq
10186 CTGTGGCTTGCTCTGCTCAACACTTGATTGAGAACTAGATGTTGATGAAG a909_ai2.seq

      TTACCTGGCTAGAATTTTTATTTTCTAAAGTAATCCCCACATCATCTGTC Majority
             10260     10270     10280     10290     10300
10239 TTACCTGGCTAGAATTTTTATTTTCTAAAGTAATCCCCACATCATCTGTC coh1_ai2.seq
10236 TTACCTGGCTAGAATTTTTATTTTCTAAAGTAATCCCCACATCATCTGTC a909_ai2.seq

      TTAGTTTCTTCAACTGTTATTGCTGGTAGAATTAAAAAATAAGTCGTTAA Majority
             10310     10320     10330     10340     10350
10289 TTAGTTTCTTCAACTGTTATTGCTGGTAGAATTAAAAAATAAGTCGTTAA coh1_ai2.seq
10286 TTAGTTTCTTCAACTGTTATTGCTGGTAGAATTAAAAAATAAGTCGTTAA a909_ai2.seq

      AAAAGTTGTTAGGATCATCAATGACCACATGATAATTTTCCACTCTTTAG Majority
             10360     10370     10380     10390     10400
10339 AAAAGTTGTTAGGATCATCAATGACCACATGATAATTTTCCACTCTTTAG coh1_ai2.seq
10336 AAAAGTTGTTAGGATCATCAATGACCACATGATAATTTTCCACTCTTTAG a909_ai2.seq
```

FIGURE 21O

```
      GGTGTTTTTTCTTTTTTAATGATTCGATTATAAAAGTTTGACACTTCTTT Majority
               10410     10420     10430     10440     10450
10389 GGTGTTTTTTCTTTTTTAATGATTCGATTATAAAAGTTTGACACTTCTTT coh1_al2.seq
10386 GGTGTTTTTTCTTTTTTAATGATTCGATTATAAAAGTTTGACACTTCTTT a909_al2.seq

      AGCATTTTTGCATCCTCCCTAACCTTAATTGATACTACTAATCTTACCTA Majority
               10460     10470     10480     10490     10500
10439 AGCATTTTTGCATCCTCCCTAACCTTAATTGATACTACTAATCTTACCTA coh1_al2.seq
10436 AGCATTTTT-CATCCTCCCTAACCTTAATTGATACTACTAATCTTACCTA a909_al2.seq

      GAGGCCATATTCTGAAAGAAATTTTACCTACAATTTGTTCTTCTGAAACA Majority
               10510     10520     10530     10540     10550
10489 GAGGCCATATTCTGAAAGAAATTTTACCTACAATTTGTTCTTCTGAAACA coh1_al2.seq
10485 GAGGCCATATTCTGAAAGAAATTTTACCTACAATTTGTTCTTCTGAAACA a909_al2.seq

      TCTCCTACAGAAGTATTTCGAGAATCAATTGAAGTTTTTCGGTTGTCTCC Majority
               10560     10570     10580     10590     10600
10539 TCTCCTACAGAAGTACTTCGAGAATCAATTGAAGTTTTTCGGTTGTCTCC coh1_al2.seq
10535 TCTCCTACAGAAGTATTTCGAGAATCAATTGAAGTTTTTCGGTTGTCTCC a909_al2.seq

      TAATACAAAAATTTTTTTATCAGGTACTTGATATGGGTATTTTATATTAC Majority
               10610     10620     10630     10640     10650
10589 TAATACAAAAATTTTTTTATCAGGTACTTGATATGGGTATTTTATATTAC coh1_al2.seq
10585 TAATACAAAAATTTTTTTATCAGGTACTTGATATGGGTATTTTATATTAC a909_al2.seq

      TATTACCGAGTGCTTTATGAATAACATATGGTTCTTTCAACTTATGTTGA Majority
               10660     10670     10680     10690     10700
10639 TATTACCGAGTGCTTTATGAATAACATATGGTTCTTTCAACTTATGTTGA coh1_al2.seq
10635 TATTACCGAGTGCTTTATGAATAACATATGGTTCTTTCAACTTATGTTGA a909_al2.seq

      TTCACGTAAACATCCCCTTGAGAATCAATATTAACCCAGTCTCCTGACTC Majority
               10710     10720     10730     10740     10750
10689 TTCACGTAAACATCCCCTTGAGAATCAATATTAACCCAGTCTCCTGACTC coh1_al2.seq
10685 TTCACGTAAACATCCCCTTGAGAATCAATATTAACCCAGTCTCCTGACTC a909_al2.seq

      TGCAATAACCCGCTTGACTAGGACCTTATTATTGTAGTAAAACGCGACAA Majority
               10760     10770     10780     10790     10800
10739 TGCAATAACCCGCTTGACTAGGACCTTATTATTGTAGTAAAACGCGACAA coh1_al2.seq
10735 TGCAATAACCCGCTTGACTAGGACCTTATTATTGTAGTAAAACGCAACAA a909_al2.seq

      CGTCTCCAGTTTTAAAATTTGAACCTTTTACTGTAAAGACTACATCACCT Majority
               10810     10820     10830     10840     10850
10789 CGTCTCCAGTTTTAAAATTTGAACCTTTTACTGTAAAGACTACATCACCT coh1_al2.seq
10785 CGTCTCCAGTTTTAAAATTTGAACCTTTTACTGTAAAGACTACATCACCT a909_al2.seq

      GCACTTAAAGTCTTATTCATTGAATGTCCGTAGATTCTTAATACAGGCAA Majority
               10860     10870     10880     10890     10900
10839 GCACTTAAAGTCTTATTCATTGAATGTCCGTAGATTCTTAATACAGGCAA coh1_al2.seq
10835 GCACTTAAAGTCTTATTCATTGAATGTCCGTAGATTCTTAATACAGGCAA a909_al2.seq

      CCATAAAACCGCAATTAAAATGGCTGTTGAGGCAACCGCCATCAAGATGT Majority
               10910     10920     10930     10940     10950
10889 CCATAAAACCGCAATTAAAATGGCTATTGAGGCAACCGCCATCAAGATGT coh1_al2.seq
10885 CCATAAAACCGCAATTAAAATGGCTGTTGAGGCAACCGCCATCAAGATGT a909_al2.seq

      ATATGGTATTTTTAATGACACTCCAAAAGCGTTTCTGATAAGTTACGCGA Majority
               10960     10970     10980     10990     11000
10939 ATATGGTATTTTTAATGACACTCCAAAAGCGTTTCTGATAAGTTACGCGA coh1_al2.seq
10935 ATATGGTATTTTTAATGACACTCCAAAAGCGTTTCTGATAAGTTACGCGA a909_al2.seq

      TCCAGTTCTTGAGATAATTTATCTGAACTAATCTGTCTTTTCATTGTCTA Majority
               11010     11020     11030     11040     11050
10989 TCCAGTTCTTGAGATAATTTATCTGAACTAATCTGTCTTTTCATTGTCTA coh1_al2.seq
10985 TCCAGTTCTTGAGATAATTTATCTGAACTAATCTGTCTTTTCATTGTCTA a909_al2.seq
```

FIGURE 21P

```
      CGCCTCTCTTATCTACTAAATTCTTTACGTTTTCTAAATAGGTATCTGCA Majority
               11060     11070     11080     11090     11100
11039 CGCCTCTCTTATCTACTAAATTCTTTACGTTTTCTAAATAGGTATCTGCA coh1_al2.seq
11035 CGCCTCTCTTATCTACTAAATTCTTTACGTTTTCTAAATAGGTATCTGCA a909_al2.seq

      GCTAACTGAGCAGTCTCAAAAATACCACTAAGCATTAAGGAAGCTTCGGC Majority
               11110     11120     11130     11140     11150
11089 GCTAACTGAGCAGTCTCAAAAATACCACTAAGCATTAAGGAAGCTTCGGC coh1_al2.seq
11085 GCTAACTGAGCAGTCTCAAAAATACCACTAAGCATTAAGGAAGCTTCGGC a909_al2.seq

      AATAGAACCAGCTTTGCTAATTTTGATTTTTTTATCATCTAGCGCTTCTT Majority
               11160     11170     11180     11190     11200
11139 AATAGAACCAGCTTTGCTAATTTTGATTTTTTTATCATCTAGCGCTTCTT coh1_al2.seq
11135 AATAGAACCAGCTTTGCTAATTTTGATTTTTTTATCATCTAGCGCTTCTT a909_al2.seq

      TAAGTTGCTGAATCTCTTTCTCTTGTTTTTCAATAAGAAGTTGCTGTTCT Majority
               11210     11220     11230     11240     11250
11189 TAAGTTGCTGAATCTCTTTCTCTTGTTTTTCAATAAGAAGTTGCTGTTCT coh1_al2.seq
11185 TAAGTTGCTGAATCTCTTTCTCTTGTTTTTCAATAAGAAGTTGCTGTTCT a909_al2.seq

      AACATAATTTCTAGCAAGTCTTTTCTTTTTAATTTTTTTAAATCTTCCAT Majority
               11260     11270     11280     11290     11300
11239 AACATAATTTCTAGCAAGTCTTTTCTTTTTAATTTTTTTAAATCTTCCAT coh1_al2.seq
11235 AACATAATTTCTAGCAAGTCTTTTCTTTTTAATTTTTTTAAATCTTCCAT a909_al2.seq

      CGCGATTACTTCCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC Majority
               11310     11320     11330     11340     11350
11289 CGCGATTACTTCCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC coh1_al2.seq
11285 CGCGATTACTTCCTTAACTGAACCTTAAATTATCGTTTAGATATTATATC a909_al2.seq

      AAAGTTCTAACCTTTAAACTCATTTTTTGTCCTGTGTTTTTTCTCAAAAA Majority
               11360     11370     11380     11390     11400
11339 AAAGTTCTAACCTTTAAACTCATTTTTTGTCCTGTGTTTTTTCTCAAAAA coh1_al2.seq
11335 AAAGTTCTAACCTTTAAACTCATTTTTTGTCCTGTGTTTTTTCTCAAAAA a909_al2.seq

      AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGAAAAATCTCATCGA Majority
               11410     11420     11430     11440     11450
11389 AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGAAAAATCTCATCGA coh1_al2.seq
11385 AGTCTATGCTAAATTAACATTTTTGATAATTTTTTGAAAAATCTCATCGA a909_al2.seq

      AGTCATTTTCTTTTTGAAAGCTCGAATTCTAGGCATTAAAAAGCCATATA Majority
               11460     11470     11480     11490     11500
11439 AGTCATTTTCTTTTTGAAAGCTCGAATTCTAGGCATTAAAAAGCCATATA coh1_al2.seq
11435 AGTCATTTTCTTTTTGAAAGCTCGAATTCTAGGCATTAAAA--------- a909_al2.seq

      TCAAATTGATATATGGCTTTTTTTATTATTTAAAACAAAAGAATCAATAG Majority
               11510     11520     11530     11540     11550
11489 TCAAATTGATATATGGCTTTTTTTATTATTTAAAACAAAAGAATCAATAG coh1_al2.seq
11476 ----------------------------------ACAAAAGAATCAATAG a909_al2.seq

      GACAATAGCGTCAATTTAGTGACATAATCTATTACAGATTAAGTTCTTTT Majority
               11560     11570     11580     11590     11600
11539 GACAATAGCGTCAATTTAGTGACATAATCTATTACAGATTAAGTTCTTTT coh1_al2.seq
11492 GACAATAGCATCAATTTAGTGACACAATCTATTACAGATTAAGTTCTTTT a909_al2.seq

      TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCATGTGAAATGTTC Majority
               11610     11620     11630     11640     11650
11589 TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCATGTGAAATGTTC coh1_al2.seq
11542 TGAATAATATAATCCAACTTTTCAACTGTTTTTTCCCAGGTGAAATGTTC a909_al2.seq

      TTTAATTCTTTTAGCAATATTCTGTTGTAGTTTCTCTCTTAATGCCTTAT Majority
               11660     11670     11680     11690     11700
11639 TTTAATTCTTTTAGCAATATTCTGTTGTAGTTTCTCTCTTAATGCCTTAT coh1_al2.seq
11592 TTTAATTCTTTTAGCAATATTCTGTTGTAATTTCTCTCTTAATGCCTTAT a909_al2.seq
```

FIGURE 21Q

```
      CTTTTACTAATAAATCAAGAGATTCATGGAGTGACTGAGTATTTTCTTCC Majority
              11710     11720     11730     11740     11750
11689 CTTTTACTAATAAATCAAGAGATTCATGGAGTGACTGAGTATTTTCTTCC coh1_a12.seq
11642 CTTTTACTAATAAATCAAGAGATTCATGGAGTGACTGAGTATTTTCTTCC a909_a12.seq

      ATGATGATTCCTAACTCAGGGCTATCAATAACTTCAACTGTTCCACCGCG Majority
              11760     11770     11780     11790     11800
11739 ATGATGATTCCTAACTCAGGGCTATCAATAACTTCAACTGTTCCACCGCG coh1_a12.seq
11692 ATGATGATTCCTAACTCAGGGCTATCAATAACTTCAACTGTTCCACCGCG a909_a12.seq

      ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAATAGAGG Majority
              11810     11820     11830     11840     11850
11789 ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAATAGAGG coh1_a12.seq
11742 ATCTGTTGCAATAATAGCACTCGAAAGTAGACCAGCTTCTAAAATAGAGG a909_a12.seq

      TTGGTAATCCCTCTGGATACATTGATGGGTAAACAAAGATATCTGTCTGT Majority
              11860     11870     11880     11890     11900
11839 TTGGTAATCCCTCTGGATACATTGATGGGTAAACAAAGATATCTGTCTGT coh1_a12.seq
11792 TTGGTAATCCCTCTGGATACATTGATGGGTAAACAAAGATATCTGTCTGT a909_a12.seq

      GCCATTAAAGACATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAAT Majority
              11910     11920     11930     11940     11950
11889 GCCATTAAAGACATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAAT coh1_a12.seq
11842 GCCATTAAAGACATAGTCTGTTCAAAGTTTAATTTCCCCAAAAAGTTAAT a909_a12.seq

      CTGTTTGGACTGATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC Majority
              11960     11970     11980     11990     12000
11939 CTGTTTGGACTGATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC coh1_a12.seq
11892 CTGTTTGGACTGATATTTCTCTTTCAAATGTGCTAATTCAGGTCCGTCTC a909_a12.seq

      CTGCAATCTGTAAATAAACATTTTCAGAGTACTGTGACATCGAAAATGCT Majority
              12010     12020     12030     12040     12050
11989 CTGCAATCTGTAAATAAACATTTTCAGAGTACTGTGACATCGAAAATGCT coh1_a12.seq
11942 CTGCAATCTGTAAATAAACATTTTCAGAGTACTGTGACATCGAAAATGCT a909_a12.seq

      TCTAAGAGCAATTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGT Majority
              12060     12070     12080     12090     12100
12039 TCTAAGAGCAATTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGT coh1_a12.seq
11992 TCTAAGAGCAATTCAATGCCTTTTTCTTTAATAATTCTACCAGCATAAGT a909_a12.seq

      GATGAAAATATCATCAGCAGATTTTTCAAGGTAAGCCGTGTCAGCAAAAT Majority
              12110     12120     12130     12140     12150
12089 GATGAAAATATCATCAGCAGATTTTTCAAGGTAAGCCGTGTCAGCAAAAT coh1_a12.seq
12042 GATGAAAATATCATCAGCAGATTTTTCAAGGTAAGCCGTACCAGCAAAAT a909_a12.seq

      CAGAGCCTAGACTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCT Majority
              12160     12170     12180     12190     12200
12139 CAGAACCTAGACTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCT coh1_a12.seq
12092 CAGAGCCTAGACTTTCAGATACCGAATTATAAATAACTCCTTTAGCTTCT a909_a12.seq

      ATATTAAAATGTTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAA Majority
              12210     12220     12230     12240     12250
12189 ATATTAAAATGTTTTAACCACTCAACGCTTCTCTTAGATACCGCATAAAA coh1_a12.seq
12142 ATATTAAAATGTTTTAACCATTCAACGCTTCTCTTGGATACCGCATAAAA a909_a12.seq

      ATCTGGACGATAGTGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAG Majority
              12260     12270     12280     12290     12300
12239 ATCTGGACGATAGTGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAG coh1_a12.seq
12192 ATCTGGACGATAGTGCTTAACACGCGCTGTGAGAAGATGTTCATAGATAG a909_a12.seq

      CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA Majority
              12310     12320     12330     12340     12350
12289 CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA coh1_a12.seq
12242 CTCCAAAGAAATCTAAAAAACGATTATTGACAGAAAAATGACTTGACCCA a909_a12.seq
```

FIGURE 21R

```
TGGTCTAAAACAATACTAGGTAAACGGTGCTTCTTTGCAAAAGATAGCCC Majority
12360 12370 12380 12390 12400
12339 TGGTCTAAAACAATACTAGGTAAACGGTGCTTCTTTGCAAAAGATAGCCC coh1_a12.seq
12292 TGGTCTAAAACAATACTAGGTAAACGGTGCTTCTTTGCAAAAGATAGCCC a909_a12.seq

TTCTAGCGTTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATAT Majority
12410 12420 12430 12440 12450
12389 TTCTAGCGTTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATAT coh1_a12.seq
12342 TTCTAGCGTTGTTAACTGAAAACGTGTATTACAAATCACAAAATCAATAT a909_a12.seq

TTTCATCTGAAACATGTTTCATCAGCGTGTTGTATTCTCGATTTTTGTTA Majority
12460 12470 12480 12490 12500
12439 TTTCATCTGAAACATGTTTCATCAGCGTGTTGTATTCTCGATTTTTGTTA coh1_a12.seq
12392 TTTCATCTGAAACATGTTTCATCAGCGTGTTGTATTCTCGATTTTTGTTA a909_a12.seq

ATAATAGGATAGCGCTGCTTGACAATGTTTTTGGTCGGTAAACGGTAAAT Majority
12510 12520 12530 12540 12550
12489 ATAATAGGATAGCGCTGCTTGACAATATTTTTGGTCGGTAAACGGTAAAT coh1_a12.seq
12442 ATAATAGGATAGCGCTGCTTGACAATGTTTTTGGTCGGTAAACGGTAAAT a909_a12.seq

TTTTCTACCCTTGTCTTCATCTATAATCGGTAAATCATCATGATTAGTTG Majority
12560 12570 12580 12590 12600
12539 TTTTCTACCCTTGTCTTCATCTATAATCGGTAAATCATCATGATTAGTTG coh1_a12.seq
12492 TTTTCTACCCTTGTCTTCATCTATAATCGGTAAATCACCATGATTAGTTG a909_a12.seq

TTACAATAACAACACGGTAGCCACGCTTAACCAAATCTGCTGTCATTTTA Majority
12610 12620 12630 12640 12650
12589 TTACAATAACAACACGGTAGCCACGCTTAACCAAATCTGCTGTCATTTTA coh1_a12.seq
12542 TTACAATAACAACACGGTAGCCACGCTTAACCAAATCTGCTGTCATTTTA a909_a12.seq

TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA Majority
12660 12670 12680 12690 12700
12639 TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA coh1_a12.seq
12592 TCTGTATAACGTTCAATACCTCCGAGGAAGGGTAGATAATATCCTGAGAA a909_a12.seq

AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTCATCAAT Majority
12710 12720 12730 12740 12750
12689 AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTCATCAAT coh1_a12.seq
12642 AACAGCAACTGTTTTTACCTTATTTTCCATATTTATCCACTTTCATCAAT a909_a12.seq

AAGCCATCTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGCTCTTTTG Majority
12760 12770 12780 12790 12800
12739 AAGCCATCTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGCTCTTTTG coh1_a12.seq
12692 AAGCCATCTTTTAAGCCTTTAATCATAGCAACTAATTTTTTGCTCTTTTG a909_a12.seq

CTCTTCTGCTACCAACACTCGAACAAATTCATTTCGCATAAATACTAAAT Majority
12810 12820 12830 12840 12850
12789 CTCTTCTGCTACCAACACTCGAACAAATTCATTTCGCATAAATACTAAAT coh1_a12.seq
12742 CTCTTCTGCTACCAACACTCGAACAAATTCATTTCGCATAAATACTAAAT a909_a12.seq

ATTTGTGCCGCTTCTTCTTACCATATTTTTTTATAATATAAATCGCATTG Majority
12860 12870 12880 12890 12900
12839 ATTTGTGCCGCTTCTTCTTACCATATTTTTTTATAATATAAATCGCATTG coh1_a12.seq
12792 ATTTGTGCCGCTTCTTCTTACCATATTTTTTTATAATATAAATCGCATTG a909_a12.seq

CGTATCATGTAATATTTTCGAAATGGTGAATGATTCAATACATGAAAAAC Majority
12910 12920 12930 12940 12950
12889 CGTATCATGTAATATTTTCGAAATGGTGAATGATTCAATACATGAAAAAC coh1_a12.seq
12842 CGTATCATGTAATATTTTCGAAATGGTGAATGATTCAATACATGAAAAAC a909_a12.seq

ATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAATTTCGTGTAAAAGAC Majority
12960 12970 12980 12990 13000
12939 ATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAATTTCGTGTAAAAGAC coh1_a12.seq
12892 ATGGCCAAATTTTTTAACTCGTGAAGAGTGTCCAATTTCGTGTAAAAGAC a909_a12.seq
```

FIGURE 21S

```
      CAATAAAATTAACCTGATAAGTCTTATATCCCATCTCTGACAGACGATAA Majority
               13010     13020     13030     13040     13050
12989 CAATAAAATTAACCTGATAAGTCTTATATCCCATCTCTGACAGACGATAA cohl_a12.seq
12942 CAATAAAATTAACCTGATAAGTCTTATATCCCATCTCTGACAGACGATAA a909_a12.seq

      TTCATTTCAGAGTCAACAAAATCAATAAACATCTCTTCTGCAAAGCCAGA Majority
               13060     13070     13080     13090     13100
13039 TTCATTTCAGAGTCAACAAAATCAATAAACATCTCTTCTGCAAAGCCAGA cohl_a12.seq
12992 TTCATTTCAGAGTCAACAAAATCAATAAACATCTCTTCTGCAAAGCCAGA a909_a12.seq

      TGTTTCTTCAAAAACGCTCGTTTTCATTAAAGCAGCCGAAGTAATACACT Majority
               13110     13120     13130     13140     13150
13089 TGTTTCTTCAAAAACGCTCGTTTTCATTAAAGCAGCCGAAGTAATACACT cohl_a12.seq
13042 TGTTTCTTCAAAAACGCTCGTTTTCATTAAAGCAGCCGAAGTAATACACT a909_a12.seq

      CTTCAATTTCTTTATAGTCAAATTCTTGCATCACTAAATCTTCACGGTTC Majority
               13160     13170     13180     13190     13200
13139 CTTCAATTTCTTTATAGTCAAATTCTTGCATAACTAAATCTTCACGGTTC cohl_a12.seq
13092 CTTCAATTTCTTTATAGTCAAATTCTTGCATCACTAAATCTTCACGGTTC a909_a12.seq

      ATATCTTGATACAAACAAGATAACATACCGACCTTAGGTAAATGAAGGTA Majority
               13210     13220     13230     13240     13250
13189 ATATCTTGATACAAACAAGATAACATACCGACCTTAGGTAAAAGAAGGTA cohl_a12.seq
13142 ATATCTTGATACAAACAAGATAACATACCGACCTTAGGTAAATGAAGGTA a909_a12.seq

      ATTTTCATAATTATCTATCAAATCACCTAGGACAACCGAATCTTGATCTA Majority
               13260     13270     13280     13290     13300
13239 ATTTTCATAATTATCTATCAAATCACCTAGGACAACCGAATCTTGATCTA cohl_a12.seq
13192 ATTTTCATAATTATCTATCAAATCACCTAGGACAACCGAATCTTGATCTA a909_a12.seq

      AAGTCAAGAACCAATCAAATTCTTGTGCTACTGCAAATTGACCGATACAG Majority
               13310     13320     13330     13340     13350
13289 AAGTCAAGAACCAATCAAATTCTTGTGCTACTGCAAATTGACCGATACAG cohl_a12.seq
13242 AAGTCAAGAACCAATCAAATTCTTGTGCTACTGCAAATTGACCAATACAG a909_a12.seq

      TTCAAAGCATATGCAATTCCTTTATTTTCTGTTAAATAATCAACAGTTAG Majority
               13360     13370     13380     13390     13400
13339 TTCAAAGCATATGCAATTCCTTTATTTTCTGTTAAATAATCAACAGTTAG cohl_a12.seq
13292 TTCAAAGCATATGCAATTCCTTTATTTTCTGTTAAATAATCAACAGTTAG a909_a12.seq

      GTGCCCCTCTTCATTATAATCGGCCACTAATTGAGAAATTTCTTCCTTAT Majority
               13410     13420     13430     13440     13450
13389 GTGCCCCTCTTCATTATAATCGGCCACTAATTGAGAAATTTCTTCCTTAT cohl_a12.seq
13342 GTGCCCCTCTTCATTATAATCGGCCACTAATTGAGAAATTTCCTCCTTAT a909_a12.seq

      TTTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT Majority
               13460     13470     13480     13490     13500
13439 TTTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT cohl_a12.seq
13392 TTTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT a909_a12.seq

      GCTCGAATGTTCTGATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAAT Majority
               13510     13520     13530     13540     13550
13489 GCTCGAATGTTCTGATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAAT cohl_a12.seq
13442 GCTCGAATGTTCTGATCTAAGCGTTCAATATTGGGGTTAAAGGTGACAAT a909_a12.seq

      ACCCGCTAAATATTTCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAA Majority
               13560     13570     13580     13590     13600
13539 ACCCGCTAAATATTTCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAA cohl_a12.seq
13492 ACCCGCTAAATATTTCATGTTCTATGCTCTTTTCTAAAATCTCTAAATAA a909_a12.seq

      CTGAATGACTGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTA Majority
               13610     13620     13630     13640     13650
13589 CTGAATGACTGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTA cohl_a12.seq
13542 CTGAATGACTGGTGCTTTGGTTATAAAAACGATACCGACATAGATAGTTA a909_a12.seq
```

FIGURE 21T

```
      CTGCTACTAAACTTTGAATGACATAATTTACCAATGATACTGACATTTGA Majority
                13660     13670     13680     13690     13700
13639 CTGCTACTAAACTTTGAATGACATAATTTACCAATGATACTGACATTTGA cohl_al2.seq
13592 CTGCTACTAAACTTTGAATGACATAATTTACCAATGATACTGACATTTGA a909_al2.seq

      GTATTGATATAATAGAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA Majority
                13710     13720     13730     13740     13750
13689 GTATTGATATAATAAAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA cohl_al2.seq
13642 GTATTGATATAATAGAGTACAGCTCCACTAAGAGTAGCAGCAATTAAATA a909_al2.seq

      GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG Majority
                13760     13770     13780     13790     13800
13739 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG cohl_al2.seq
13692 GCGCAGCATTCCTCTTGTTAATTCTTTAAAAGTAAATACATCTCTTAAAG a909_al2.seq

      AGATAGCTTGATATAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA Majority
                13810     13820     13830     13840     13850
13789 AGATAGCTTGATAGAGGGAGACAATAAATTCAGTAATAACTGTAGAGATA cohl_al2.seq
13742 AGATAGCTTGATATAGAGAGACAATAAATTCAGTAATAACTGTAGAGATA a909_al2.seq

      ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC Majority
                13860     13870     13880     13890     13900
13839 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC cohl_al2.seq
13792 ATAGCTCCCATAGCACCTAAAATTGGTATTAAAAGTATATTAAGCACAAC a909_al2.seq

      ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC Majority
                13910     13920     13930     13940     13950
13889 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC cohl_al2.seq
13842 ATTTGCCACAAGTCCAATAACTGCAGACATTGTGTAAGCTTTTGTACGTC a909_al2.seq

      TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA Majority
                13960     13970     13980     13990     14000
13939 TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA cohl_al2.seq
13892 TTGAAGCCAGTAGATACTGTGTCCCTAAAGCGTTACCATAAGAAATGCAA a909_al2.seq

      ATGATCATCAAA                                       Majority
                14010
13989 ATGATCATCAAA                                       cohl_al2.seq
13942 ATGATCATCAAA                                       a909_al2.seq
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 21U

```
    MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV Majority
             10        20        30        40        50
1   MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV sag645_2603.pep
1   MKLSKKLLFSAAVLTIVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV sag645_a909.pep
1   MKLSKKLLFSAAVLTIVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV sag645_cjb111.pep
1   MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV sag645_coh1.pep
1   MKLSKKLLFSAAVLTMVAGSTVEPVAQFATGMSIVRAAEVSQERPAKTTV sag645_nem316.pep

    NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK Majority
             60        70        80        90        100
51  NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK sag645_2603.pep
51  NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK sag645_a909.pep
51  NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK sag645_cjb111.pep
51  NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK sag645_coh1.pep
51  NIYKLQADSYKSEITSNGGIENKDGEVISNYAKLGDNVKGLQGVQFKRYK sag645_nem316.pep

    VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK Majority
             110       120       130       140       150
01  VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK sag645_2603.pep
01  VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK sag645_a909.pep
01  VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK sag645_cjb111.pep
01  VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK sag645_coh1.pep
01  VKTDISVDELKKLTTVEAADAKVGTILEEGVSLPQKTNAQGLVVDALDSK sag645_nem316.pep

    SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN Majority
             160       170       180       190       200
51  SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN sag645_2603.pep
51  SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN sag645_a909.pep
51  SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN sag645_cjb111.pep
51  SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN sag645_coh1.pep
51  SNVRYLYVEDLKNSPSNITKAYAVPFVLELPVANSTGTGFLSEINIYPKN sag645_nem316.pep

    VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD Majority
             210       220       230       240       250
01  VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD sag645_2603.pep
01  VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD sag645_a909.pep
01  VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD sag645_cjb111.pep
01  VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD sag645_coh1.pep
01  VVTDEPKTDKDVKKLGQDDAGYTIGEEFKWFLKSTIPANLGDYEKFEITD sag645_nem316.pep

    KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK Majority
             260       270       280       290       300
51  KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK sag645_2603.pep
51  KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK sag645_a909.pep
51  KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK sag645_cjb111.pep
51  KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK sag645_coh1.pep
51  KFADGLTYKSVGKIKIGSKTLNRDEHYTIDEPTVDNQNTLKITFKPEKFK sag645_nem316.pep

    EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE Majority
             310       320       330       340       350
01  EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE sag645_2603.pep
01  EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE sag645_a909.pep
01  EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE sag645_cjb111.pep
01  EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE sag645_coh1.pep
01  EIAELLKGMTLVKNQDALDKATANTDDAAFLEIPVASTINEKAVLGKAIE sag645_nem316.pep

    NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE Majority
             360       370       380       390       400
51  NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE sag645_2603.pep
51  NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE sag645_a909.pep
51  NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE sag645_cjb111.pep
51  NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE sag645_coh1.pep
51  NTFELQYDHTPDKADNPKPSNPPRKPEVHTGGKRFVKKDSTETQTLGGAE sag645_nem316.pep
```

Figure 22

```
    F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  Majority
                    410                 420                 430                 440                 450
401 F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  sag645_2603.pep
401 F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  sag645_a909.pep
401 F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  sag645_cjb111.pep
401 F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  sag645_coh1.pep
401 F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K  sag645_nem316.pep

    G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  Majority
                    460                 470                 480                 490                 500
451 G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  sag645_2603.pep
451 G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  sag645_a909.pep
451 G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  sag645_cjb111.pep
451 G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  sag645_coh1.pep
451 G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D  sag645_nem316.pep

    I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  Majority
                    510                 520                 530                 540                 550
501 I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  sag645_2603.pep
501 I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  sag645_a909.pep
501 I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  sag645_cjb111.pep
501 I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  sag645_coh1.pep
501 I T V D S A D A T P D T I K N N K R P S I P N T G G I G T A I F V A I G A A V M A F A V K G M K R R  sag645_nem316.pep

    T K D N                                                                                            Majority

551 T K D N                                                                                            sag645_2603.pep
551 T K D N                                                                                            sag645_a909.pep
551 T K D N                                                                                            sag645_cjb111.pep
551 T K D N                                                                                            sag645_coh1.pep
551 T K D N                                                                                            sag645_nem316.pep
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 22A

```
    M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  Majority
                      10                  20                  30                  40                  50
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_2603.pep
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_coh1.pep
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_cjb111.pep
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_18rs21.pep
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_nem316.pep
1   M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A  sag649_cjb111.pep

    T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  Majority
                      60                  70                  80                  90                  100
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_2603.pep
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_coh1.pep
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_cjb111.pep
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_18rs21.pep
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_nem316.pep
51  T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G  sag649_cjb111.pep

    Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  Majority
                      110                 120                 130                 140                 150
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_2603.pep
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_coh1.pep
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_cjb111.pep
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_18rs21.pep
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_nem316.pep
101 Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E  sag649_cjb111.pep

    N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  Majority
                      160                 170                 180                 190                 200
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_2603.pep
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_coh1.pep
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_cjb111.pep
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_18rs21.pep
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_nem316.pep
151 N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K K I T G V N D L D K N K  sag649_cjb111.pep

    Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  Majority
                      210                 220                 230                 240                 250
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_2603.pep
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_coh1.pep
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_cjb111.pep
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_18rs21.pep
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_nem316.pep
201 Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E  sag649_cjb111.pep

    A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  Majority
                      260                 270                 280                 290                 300
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_2603.pep
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_coh1.pep
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_cjb111.pep
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_18rs21.pep
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_nem316.pep
251 A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V  sag649_cjb111.pep

    S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  Majority
                      310                 320                 330                 340                 350
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_2603.pep
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_coh1.pep
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_cjb111.pep
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_18rs21.pep
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_nem316.pep
301 S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q  sag649_cjb111.pep

    F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  Majority
                      360                 370                 380                 390                 400
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_2603.pep
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_coh1.pep
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_cjb111.pep
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_18rs21.pep
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_nem316.pep
351 F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T M S Y A I N F N P Y I S  sag649_cjb111.pep
```

Figure 23

```
    T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V Majority
                      410                 420                 430                 440                 450
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_2603.pep
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_coh1.pep
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_cjb111.pep
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_18rs21.pep
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_nem316.pep
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_cjb111.pep

    P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T Majority
                      460                 470                 480                 490                 500
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_2603.pep
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_coh1.pep
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_cjb111.pep
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_18rs21.pep
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_nem316.pep
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I Y L Y W R D Y N W V Y P F D P K T sag649_cjb111.pep

    K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E Majority
                      510                 520                 530                 540                 550
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_2603.pep
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_coh1.pep
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_cjb111.pep
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_18rs21.pep
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_nem316.pep
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_cjb111.pep

    K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G Majority
                      560                 570                 580                 590                 600
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_2603.pep
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_coh1.pep
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_cjb111.pep
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_18rs21.pep
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_nem316.pep
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_cjb111.pep

    E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T Majority
                      610                 620                 630                 640                 650
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_2603.pep
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_coh1.pep
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_cjb111.pep
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_18rs21.pep
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_nem316.pep
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_cjb111.pep

    Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K Majority
                      660                 670                 680                 690                 700
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_2603.pep
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_coh1.pep
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_cjb111.pep
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_18rs21.pep
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_nem316.pep
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_cjb111.pep

    S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L Majority
                      710                 720                 730                 740                 750
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_2603.pep
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_coh1.pep
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_cjb111.pep
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_18rs21.pep
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_nem316.pep
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_cjb111.pep

    L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S Majority
                      760                 770                 780                 790                 800
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_2603.pep
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_coh1.pep
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_cjb111.pep
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_18rs21.pep
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_nem316.pep
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_cjb111.pep
```

FIGURE 23A

```
    S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  Majority
                        810                 820                 830                 840                 850
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_2603.pep
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_cohl.pep
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_cjb111.pep
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_18rs21.pep
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_nem316.pep
801 S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N  sag649_cjb111.pep

    T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  Majority
                        860                 870                 880                 890
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_2603.pep
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_cohl.pep
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_cjb111.pep
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_18rs21.pep
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_nem316.pep
851 T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L  sag649_cjb111.pep
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 23B

```
    MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ Majority
             10        20        30        40        50
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_2603.pep
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_515.pep
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_18rs21.seq
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_cjb111.pep
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_h36b.pep
1   MRKYQKFSKILTLSLFCLSQIPLNTNVLGESTVPENGAKGKLVVKKTDDQ sag1408_nem316.pep

    NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP Majority
             60        70        80        90       100
51  NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP sag1408_2603.pep
51  NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP sag1408_515.pep
51  NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP sag1408_18rs21.seq
51  NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP sag1408_cjb111.pep
51  NKPLSKATFVLKPTSHSESKVEKVTIEVTGEATFDNLIPGDYTLSEETAP sag1408_h36b.pep
51  NKPLSKATFVLKTTAHPESKIEKVTAELTGEATFDNLIPGDYTLSEETAP sag1408_nem316.pep

    EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYE Majority
            110       120       130       140       150
101 EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYE sag1408_2603.pep
101 EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYE sag1408_515.pep
101 EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYE sag1408_18rs21.seq
101 EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNQEELDKQYPPTGIYE sag1408_cjb111.pep
101 EGYKKTIQTWQVKVESNGKTTIQNSDDKKSIIEQKQEELDKQYPITGAYE sag1408_h36b.pep
101 EGYKKTNQTWQVKVESNGKTTIQNSGDKNSTIGQNHEELDKQYPPTGIYE sag1408_nem316.pep

    DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV Majority
            160       170       180       190       200
151 DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV sag1408_2603.pep
151 DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV sag1408_515.pep
151 DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV sag1408_18rs21.seq
151 DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV sag1408_cjb111.pep
151 DTKESYALEHVKNSIPNGKLEAKAVNPYSSEGEHIREIQEGTLSKRISEV sag1408_h36b.pep
151 DTKESYKLEHVKGSVPNGKSEAKAVNPYSSEGEHIREIPEGTLSKRISEV sag1408_nem316.pep

    GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR Majority
            210       220       230       240       250
201 GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR sag1408_2603.pep
201 GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR sag1408_515.pep
201 GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR sag1408_18rs21.seq
201 GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR sag1408_cjb111.pep
201 NDLDHNKYKIELTVSGKSIIKIINKDFPLDVVFVLDNSNSMKNNG----K sag1408_h36b.pep
201 GDLAHNKYKIELTVSGKTIVKPVDKQKPLDVVFVLDNSNSMNNDGPNFQR sag1408_nem316.pep

    HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE Majority
            260       270       280       290       300
251 HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE sag1408_2603.pep
251 HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE sag1408_515.pep
251 HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE sag1408_18rs21.seq
251 HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE sag1408_cjb111.pep
247 NNKAKKAGEAVETIIKDVLGANVINRAALVTYGSDIFDGRIVKVIKGFKE sag1408_h36b.pep
251 HNKAKKAAEALGTAVKDILGANSDNRVALVTYGSDIFDGRSVDVVKGFKE sag1408_nem316.pep

    DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTNGL Majority
            310       320       330       340       350
301 DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTNGL sag1408_2603.pep
301 DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTNGL sag1408_515.pep
301 DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTYGL sag1408_18rs21.seq
301 DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPKAKWGSTTNGL sag1408_cjb111.pep
297 DP-YYGLETSFTVQTNDYSYKKFTNIAADIIKKIPKEAPIAKWGGTSIGL sag1408_h36b.pep
301 DDKYYGLQTKFTIQTENYSHKQLTNNAEEIIKRIPTEAPRAKWGSTTNGL sag1408_nem316.pep

    TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTR Majority
            360       370       380       390       400
351 TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTR sag1408_2603.pep
351 TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTR sag1408_515.pep
351 TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTR sag1408_18rs21.seq
351 TPEQQKEYYLSKVGETFTMKAFMEADDILSQVNRNSQKIIVHVTDGVPTR sag1408_cjb111.pep
346 TPEKKREYDLSKVGETFTMKAFMEADTILSSTQRKSRKIIVHITDGVPTR sag1408_h36b.pep
351 TPEQQKQYYLSKVGETFTMKAFMEADDILSQVDRNSQKIIVHITDGVPTR sag1408_nem316.pep
```

Figure 24

```
    S Y A I N N F K L G A S Y E S Q F E Q M K K N G Y L N K S N F L L T D K P E D I K G N G E S Y F L F  Majority
                      410                 420                 430                 440                 450
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G Y L N K S N F L L T D K P E D I K G N G E S Y F L F  sag1408_2603.pep
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G Y L N K S N F L L T D K P D D I K G N G E S Y F L F  sag1408_515.pep
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G Y L N K S N F L L T D K P E D I K G N G E S Y F L F  sag1408_18rs21.seq
396 S Y A I N S F V K G S T Y A N Q F E R I K F K G Y L D K A N Y T I T D D P E K I K G N G E S Y F L F  sag1408_cjb111.pep
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G Y L N K S N F L L T D K P E D I K G N G E S Y F L F  sag1408_h36b.pep
                                                                                                        sag1408_nem316.pep

    P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S  Majority
                      460                 470                 480                 490                 500
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S  sag1408_2603.pep
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S  sag1408_515.pep
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S  sag1408_18rs21.seq
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V R E H G T P T K L Y I N S  sag1408_cjb111.pep
446 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V R E H G T P T K L Y I N S  sag1408_h36b.pep
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V R E H G T P T K L Y I N S  sag1408_nem316.pep

    L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T  Majority
                      510                 520                 530                 540                 550
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E H Y K K N Q D G T F Q K L K E E A F K L S D G E I T  sag1408_2603.pep
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E H Y K K N Q D G T F Q K L K E E A F K L S D G E I T  sag1408_515.pep
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E H Y K K N Q D G T F Q K L K E E A F K L S D G E I T  sag1408_18rs21.seq
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T  sag1408_cjb111.pep
496 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T  sag1408_h36b.pep
501 L K Q K N Y D I F N F G I D I S A F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T  sag1408_nem316.pep

    E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I  Majority
                      560                 570                 580                 590                 600
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I  sag1408_2603.pep
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I  sag1408_515.pep
551 E L M K S F S S K P E Y Y T P I V T S S D A S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I  sag1408_18rs21.seq
546 E L M N S F S S K P E Y Y T P I V T S A D V S N N E I L S K I Q Q Q F E K I L T K E N S I V N G T I  sag1408_cjb111.pep
551 E L M K S F S S K P E Y Y T P I V T S S D A S N N E I L S K I Q Q Q F E K V L T K E N S I V N G T I  sag1408_h36b.pep
                                                                                                        sag1408_nem316.pep

    E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K  Majority
                      610                 620                 630                 640                 650
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K  sag1408_2603.pep
601 E D P M G D K I N L Q L G N G Q I L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K  sag1408_515.pep
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K  sag1408_18rs21.seq
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S I M K D S I A T G G P N N D G G I L K  sag1408_cjb111.pep
596 E D P M G D K I N L H L G N G Q T L Q P S D Y T L Q G N D G S I M K D S I A T G G P N N D G G I L K  sag1408_h36b.pep
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S I M K D S I A T G G P N N D G G I L K  sag1408_nem316.pep

    G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  Majority
                      660                 670                 680                 690                 700
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_2603.pep
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_515.pep
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_18rs21.seq
651 G V K L E Y I K N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_cjb111.pep
646 G V K L E Y I K N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_h36b.pep
651 G V K L E Y I K N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L  sag1408_nem316.pep

    N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  Majority
                      710                 720                 730                 740                 750
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_2603.pep
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_515.pep
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_18rs21.seq
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_cjb111.pep
696 N P K S E I P D T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_h36b.pep
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L  sag1408_nem316.pep

    L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  Majority
                      760                 770                 780                 790                 800
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_2603.pep
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_515.pep
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_18rs21.seq
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_cjb111.pep
746 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_h36b.pep
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I  sag1408_nem316.pep
```

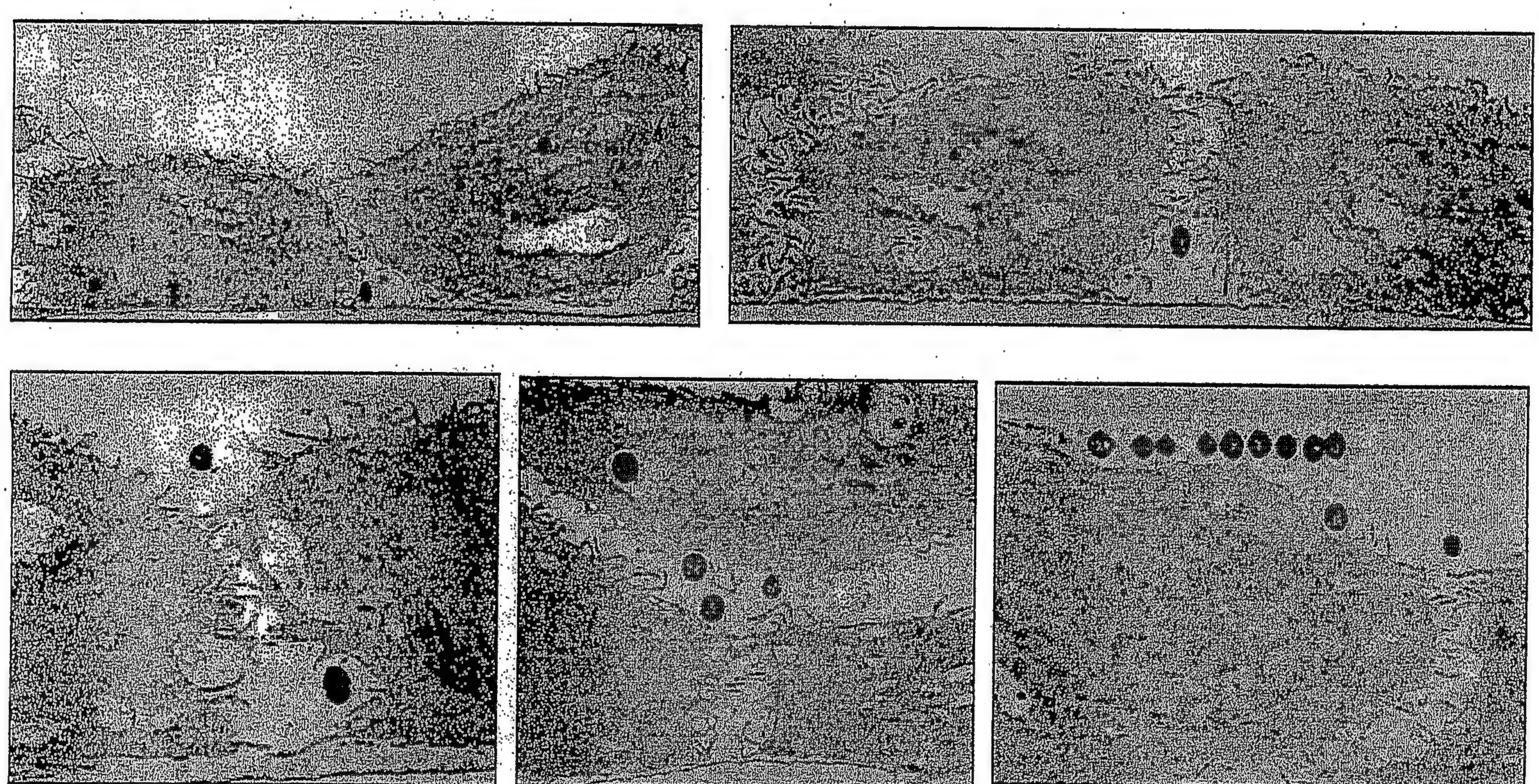
FIGURE 24A

```
     E A V S P E D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N  Majority
                       810                 820                 830                 840                 850
801  E A V S P E D Y Q K I T N K P I L T F E V V K G S I K N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_2603.pep
801  E A V S P E D Y Q K I T N K P I L T F E V V K G S I K N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_515.pep
801  E A V S P E D Y Q K I T N K P I L T F E V V K G S I K N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_18rs21.seq
801  E A V S P K D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_cjb111.pep
796  E A V S P K D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_h36b.pep
801  E A V S P K D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N  sag1408_nem316.pep

     T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K  Majority
                       860                 870                 880                 890                 900
851  T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I K K  sag1408_2603.pep
851  T H I P P K G I I P K T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I K K  sag1408_515.pep
851  T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I K K  sag1408_18rs21.seq
851  T H I P P K G I I P M T G G K G I L S F I L I G G S M M S I A G G I Y I W K R Y K K S S D I S R E K  sag1408_cjb111.pep
846  T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R H K K S S D A S I E K  sag1408_h36b.pep
851  T H I P P K G I I P M T G G K G I L S F I L I G G S M M S I A G G I Y I W K R Y K K S S D I S R E K  sag1408_nem316.pep

     D                                                                                                  Majority

901  D                                                                                                  sag1408_2603.pep
901  D                                                                                                  sag1408_515.pep
901  D                                                                                                  sag1408_18rs21.seq
901  D                                                                                                  sag1408_cjb111.pep
896  D                                                                                                  sag1408_h36b.pep
901  D                                                                                                  sag1408_nem316.pep
```

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 24B

Figure 25: GBS closely associate with tight junctions and cross the monolayer by a paracellular route



Transmission Electron Microscopy images of GBS infection of ME180 cervical epithelial cells.

Figure 26: GBS infection of ME180 cells

GBS infection of ME180 cells

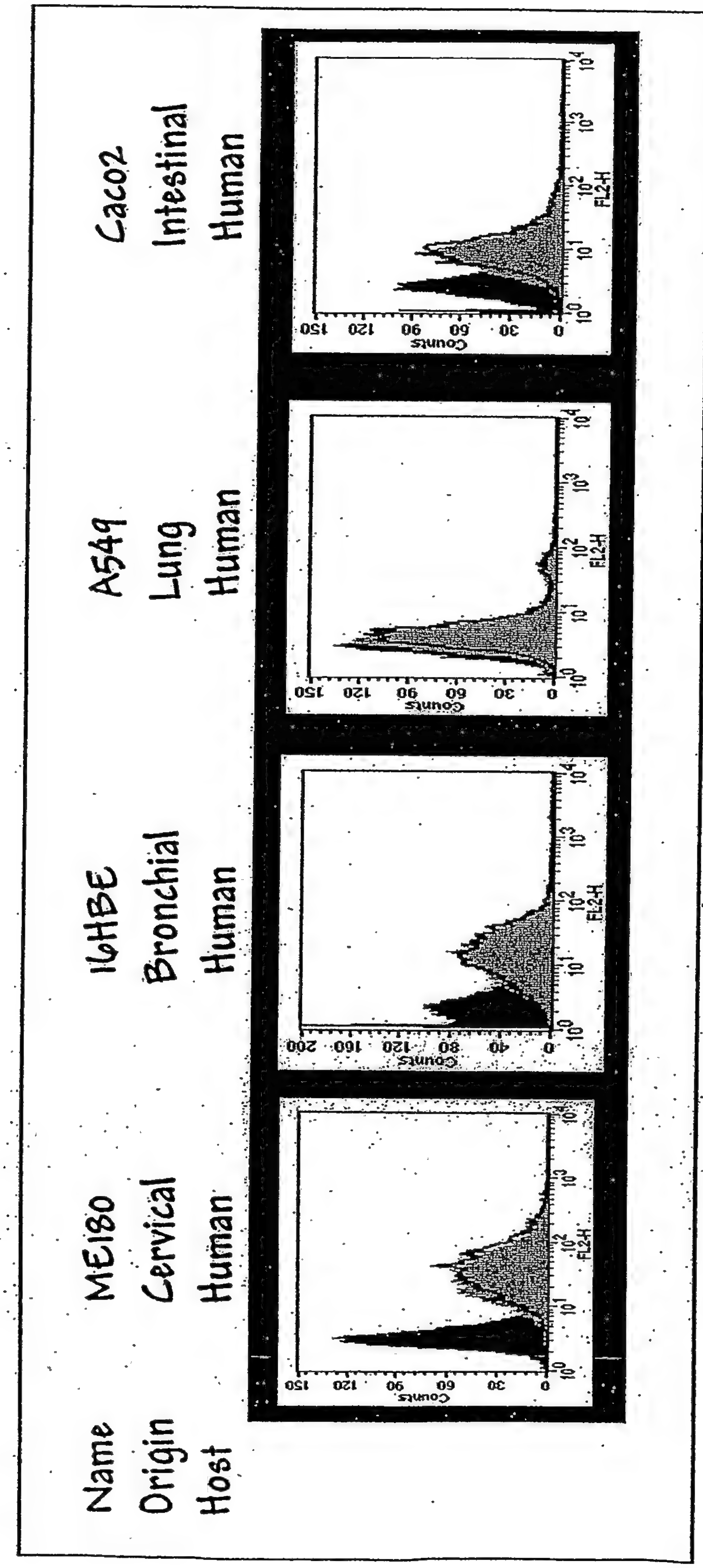

Name
Origin
Host
ME180
Cervical
Human
16HBE
Bronchial
Human
A549
Lung
Human
Caco2
Intestinal
Human
Counts
FL2-H

Figure 27

Figure 28

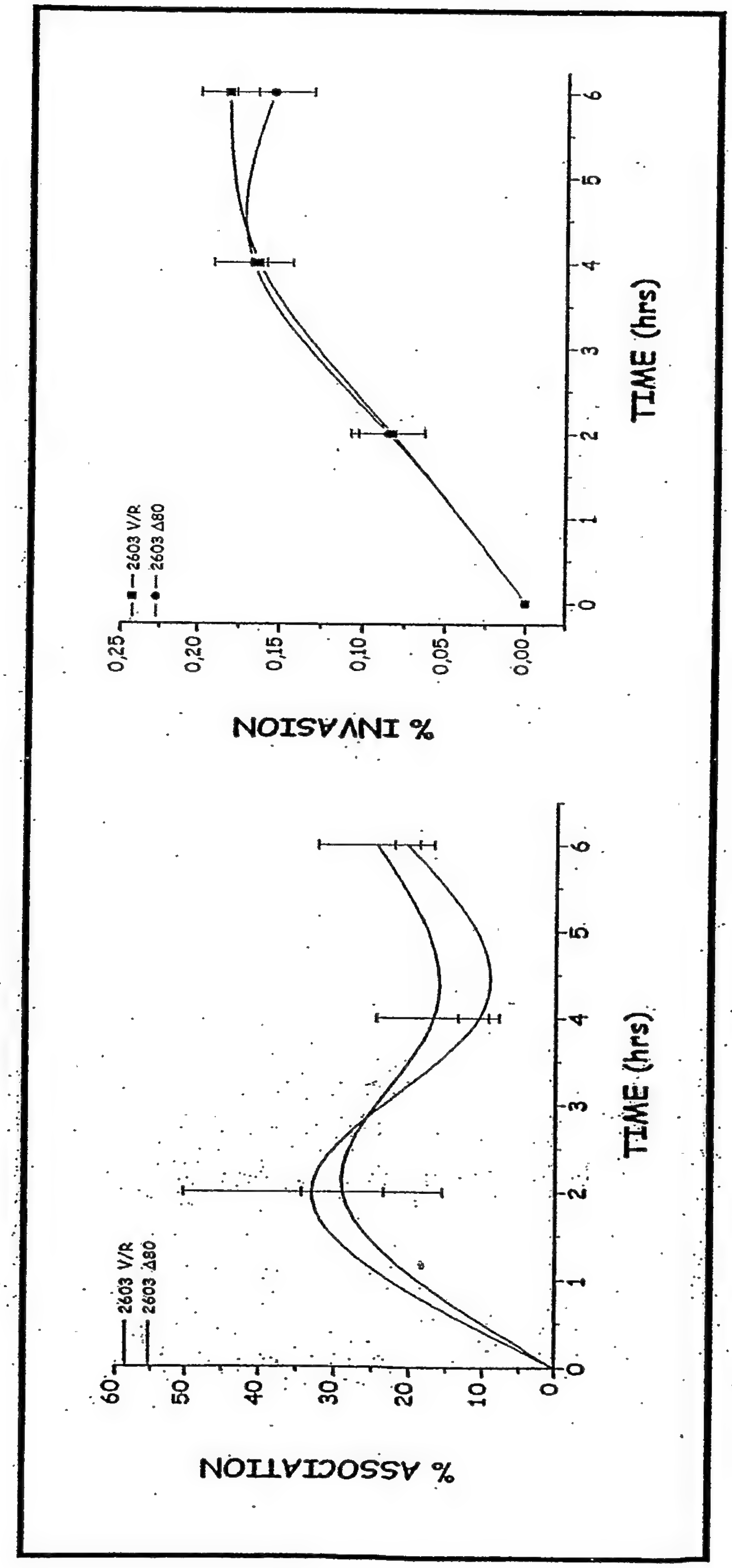

% ASSOCIATION
2603 V/R
2603 Δ80
TIME (hrs)
% INVASION
2603 V/R
2603 Δ80
TIME (hrs)

Figure 29

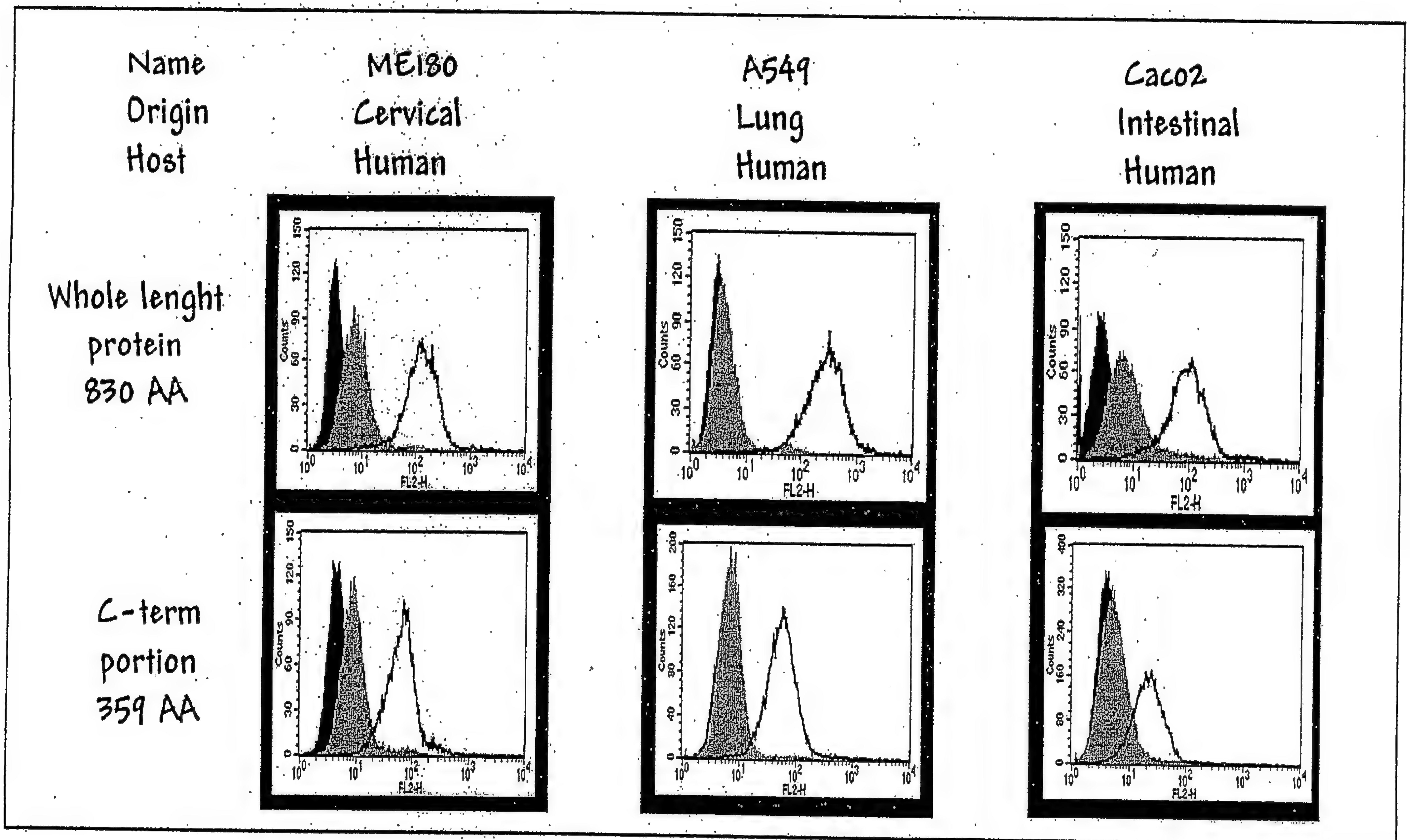
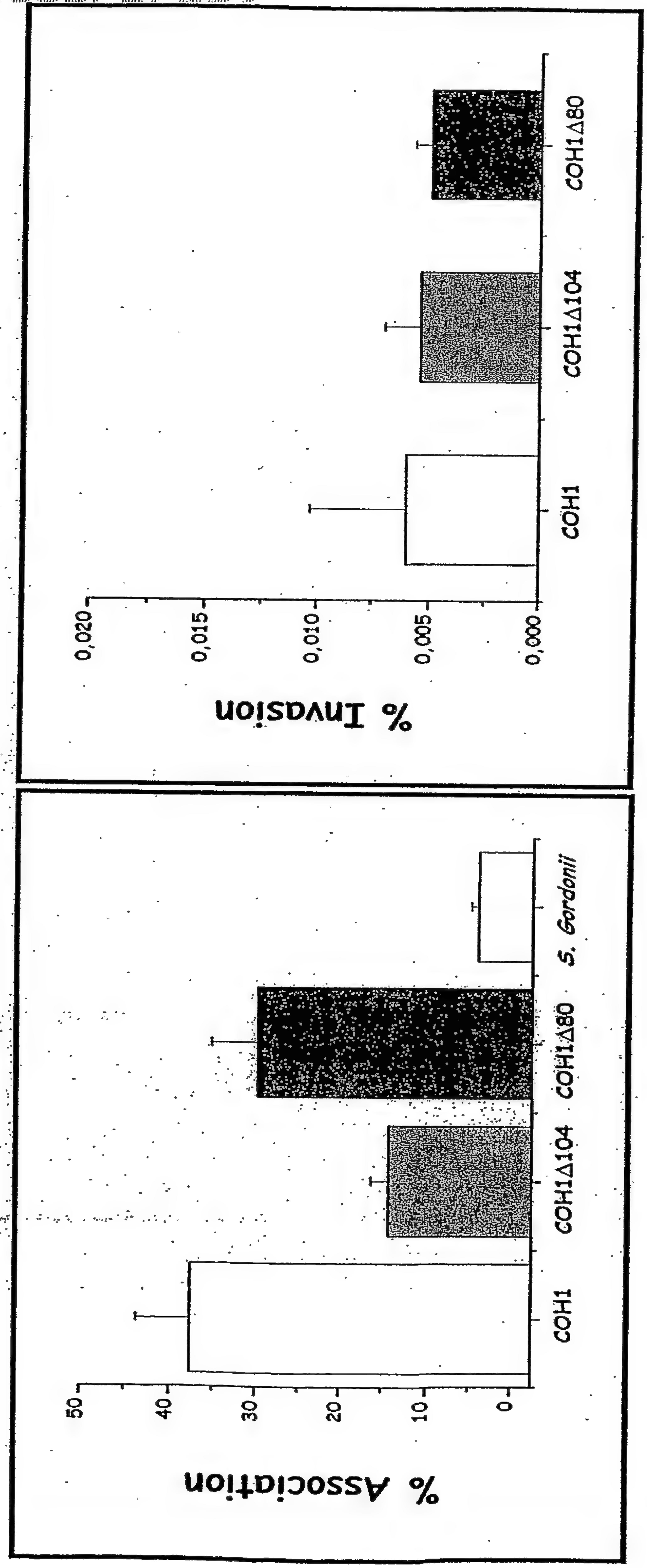

Name
Origin
Host
ME180
Cervical
Human
A549
Lung
Human
Caco2
Intestinal
Human
Whole lenght
protein
830 AA
C-term
portion
359 AA
Counts
FL2-H

Figure 30


% Invasion
0,020
0,015
0,010
0,005
0,000
COH1
COH1Δ104
COH1Δ80
% Association
50
40
30
20
10
0
COH1
COH1Δ104
COH1Δ80
S. Gordonii

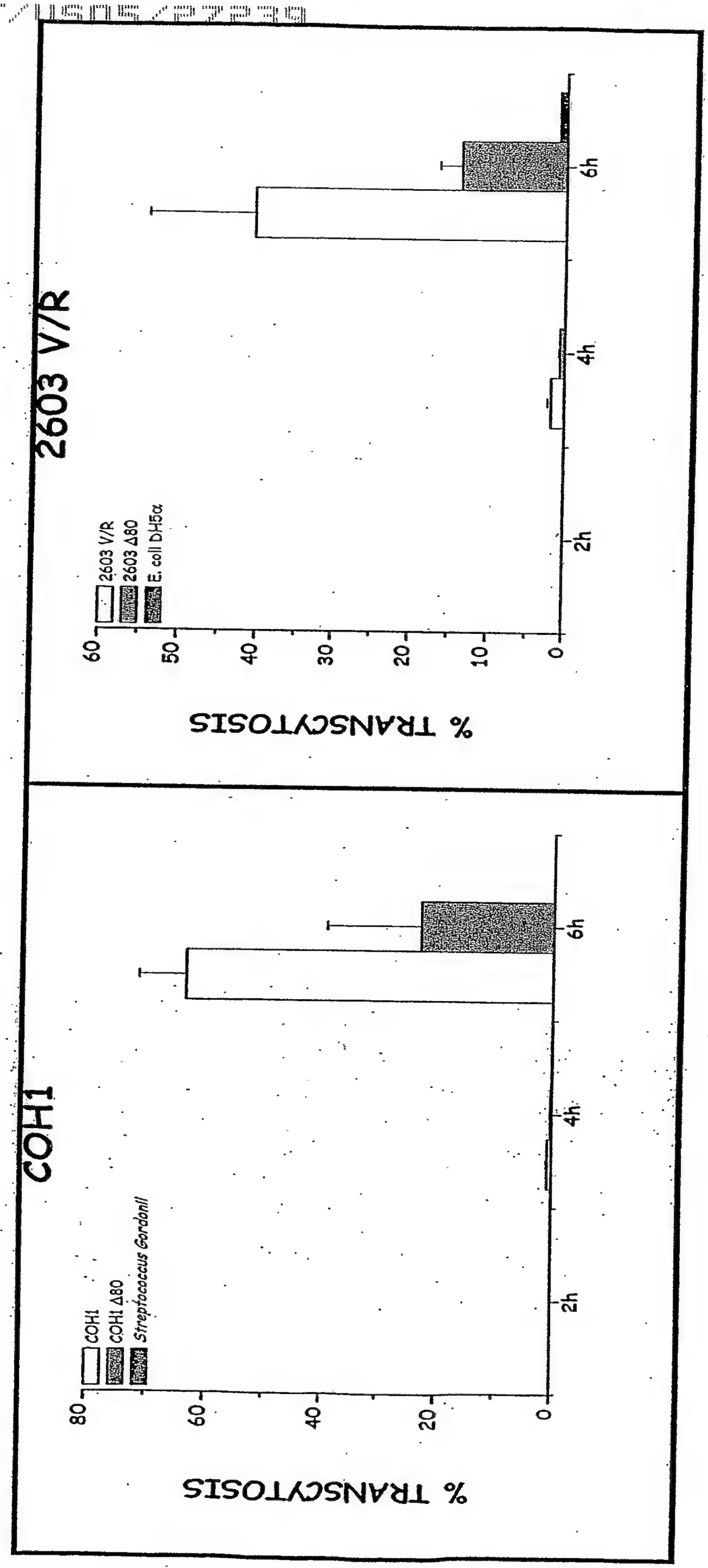
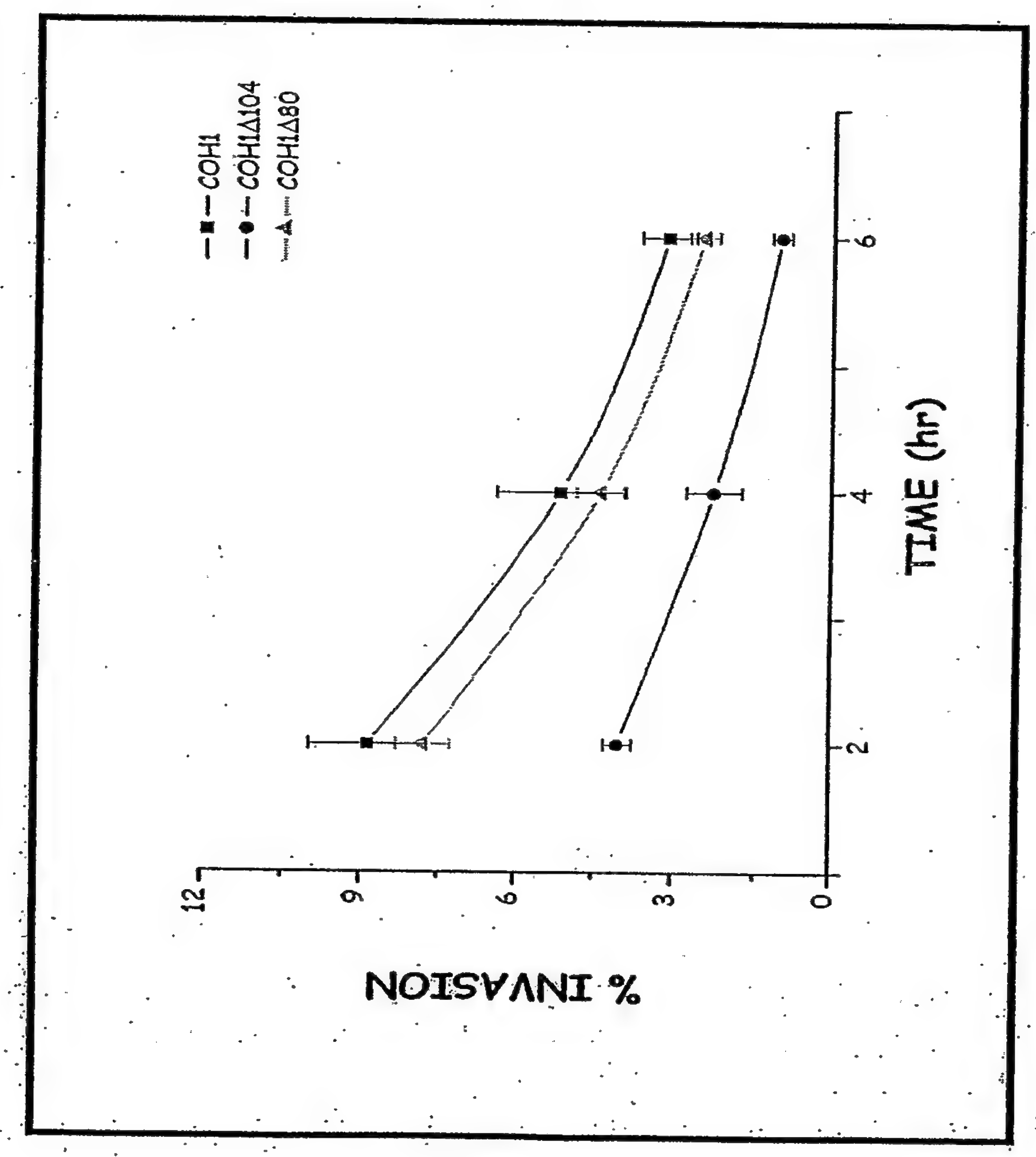

2603 V/R
% TRANSCYTOSIS
0
10
20
30
40
50
60
2603 V/R
2603 Δ80
E. coli DH5α
2h
4h
6h
COH1
% TRANSCYTOSIS
0
20
40
60
80
COH1
COH1 Δ80
Streptococcus Gordonii
2h
4h
6h

Figure 31

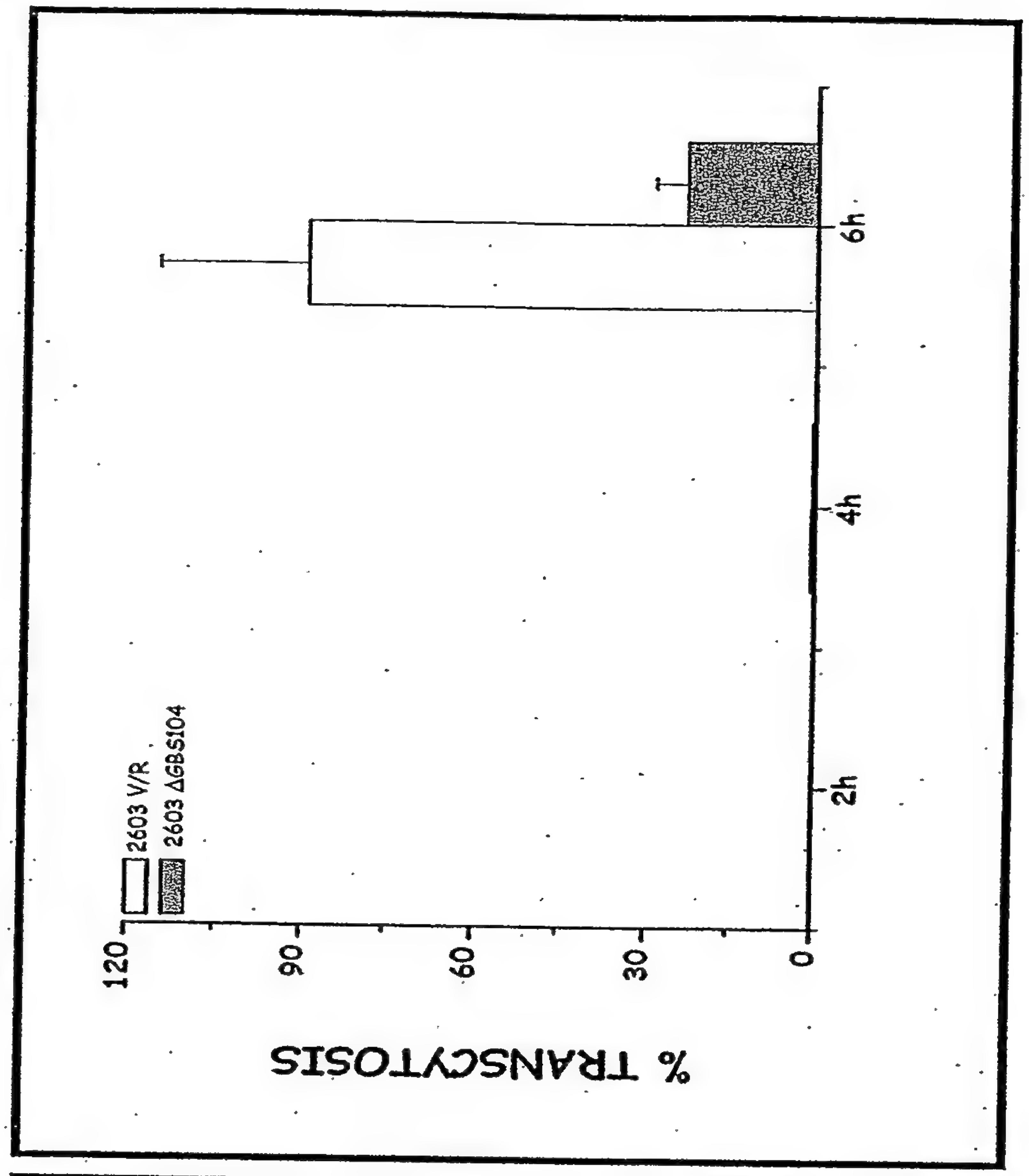
Figure 32



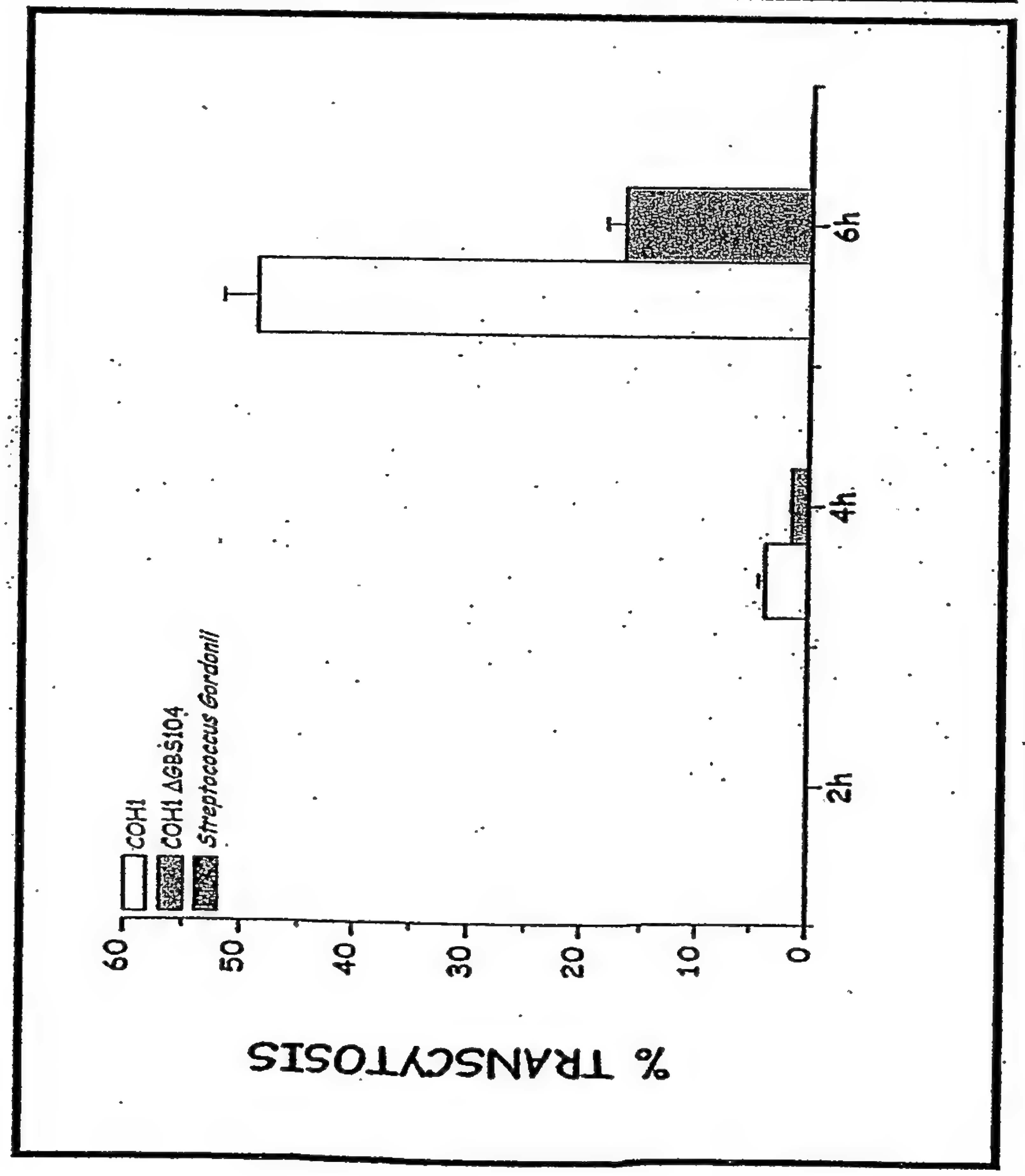
PCT/US05/27239

Figure 33

2603 V/R



COH1



GBS STRAIN COH1 over GBS80

Negative staining EM

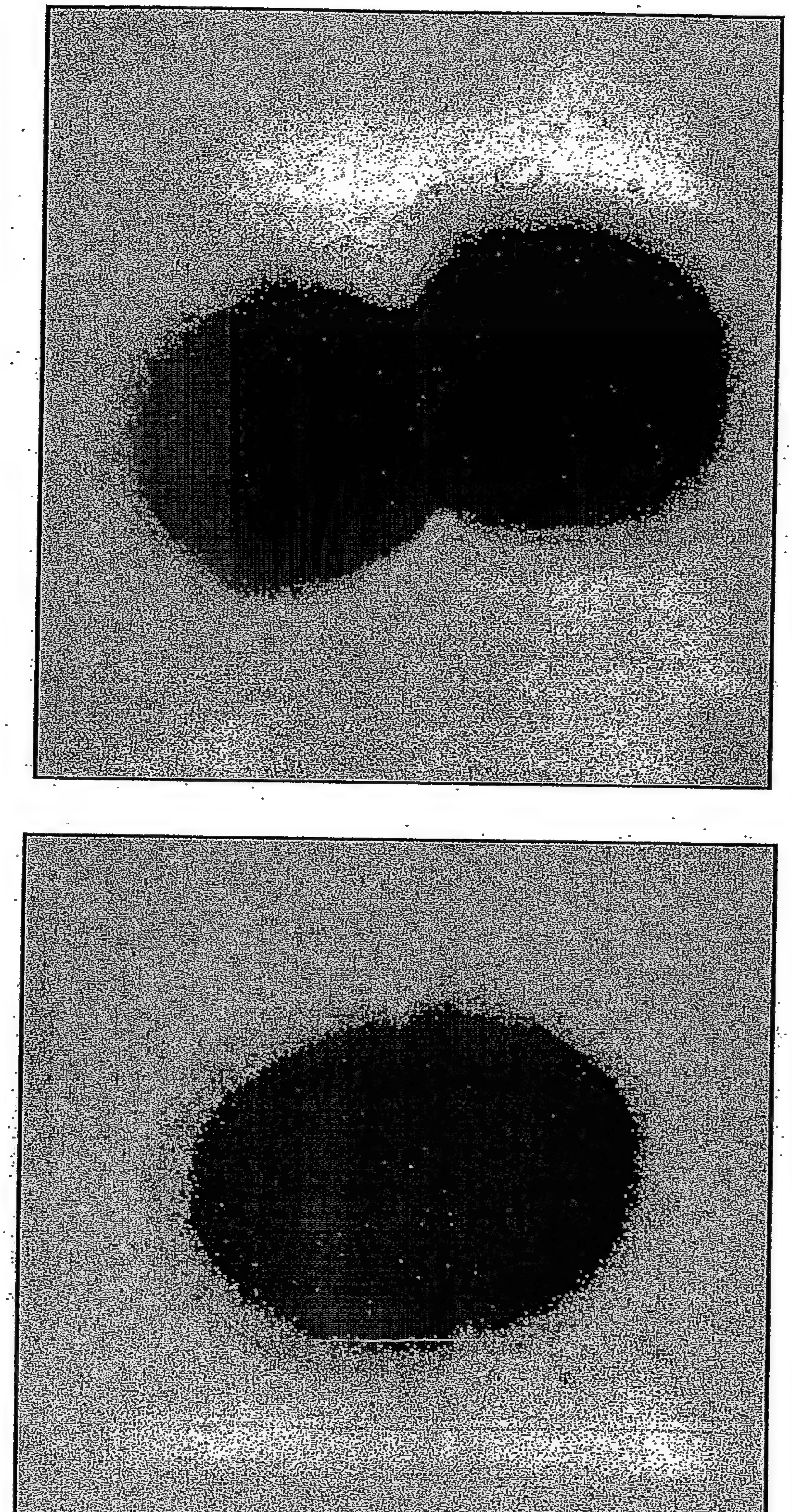


Figure 34

Figure 35

GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 10nm)

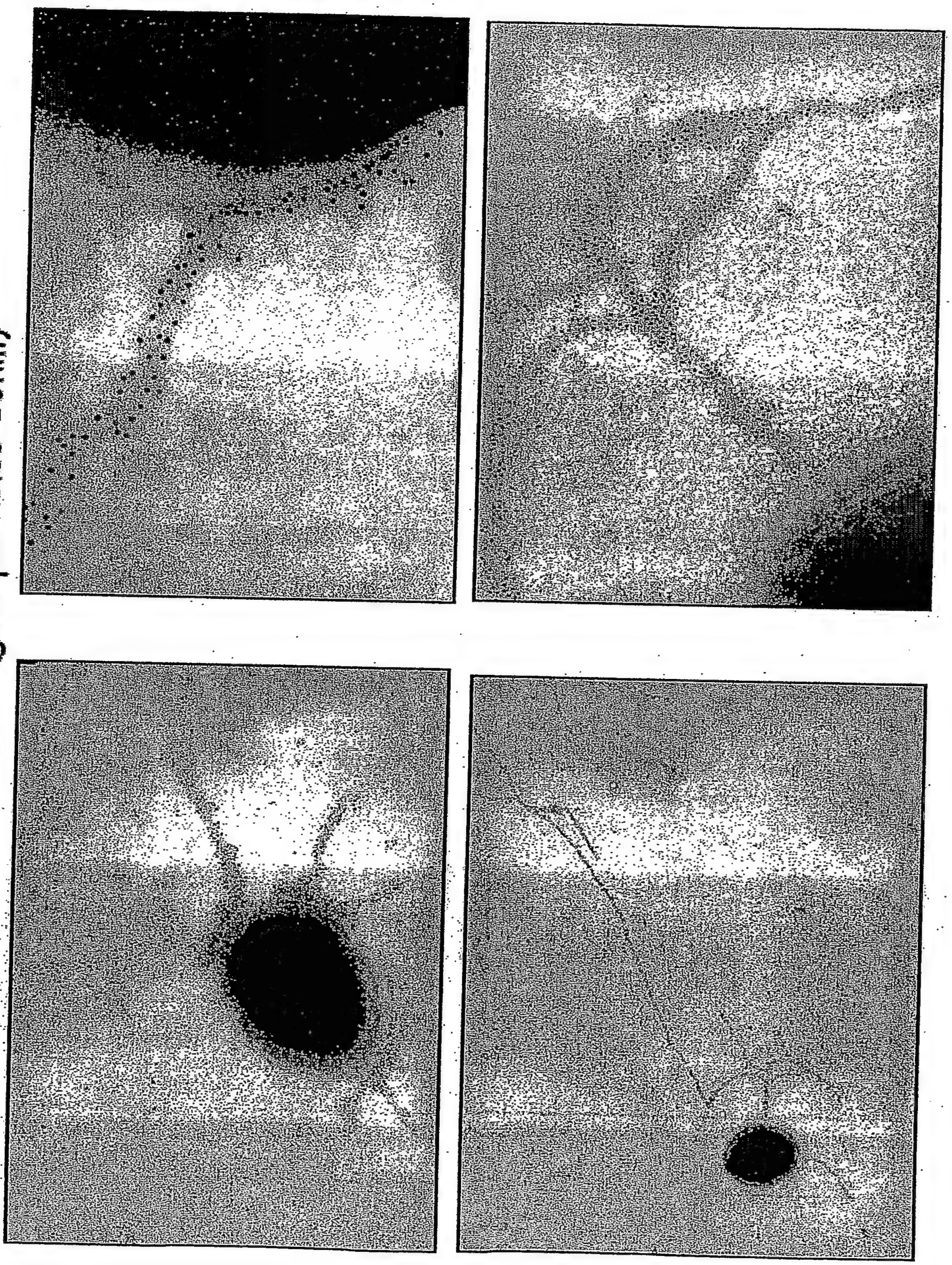


Figure 36

GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 10nm)

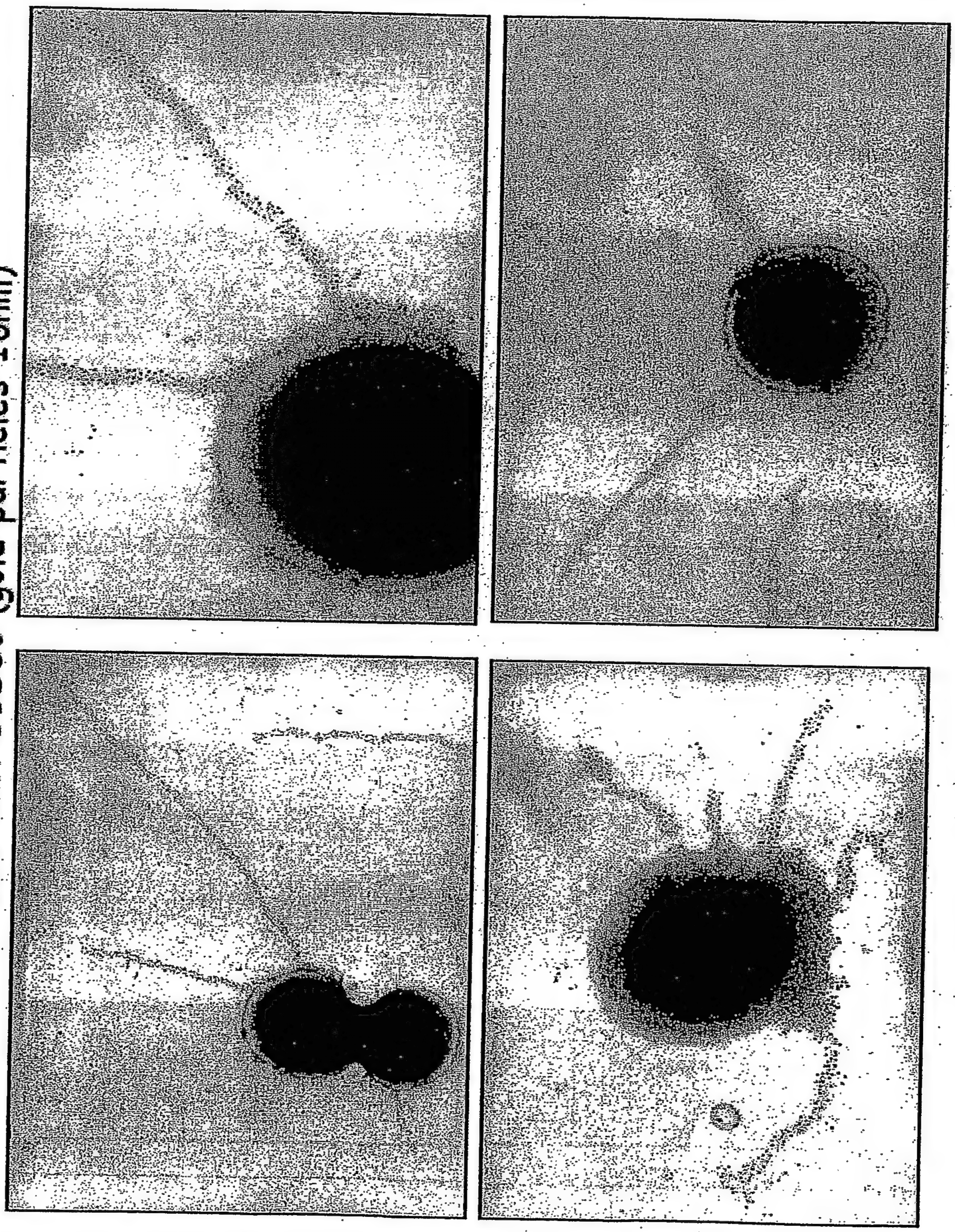


Figure 37

GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 20nm)

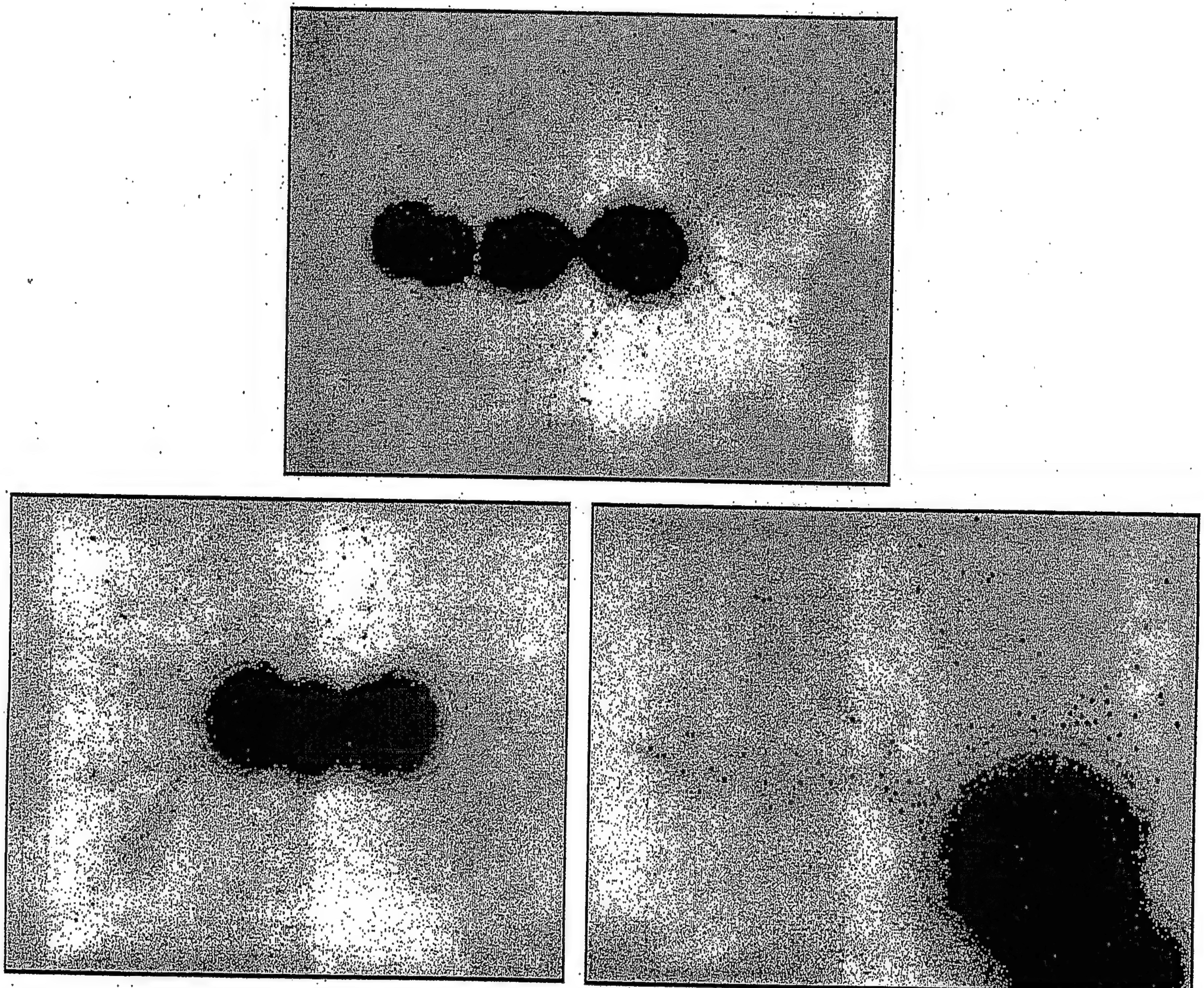


Figure 38

GBS STRAIN COH1 over GBS80

IEM anti-GBS104 (gold particles 10nm)

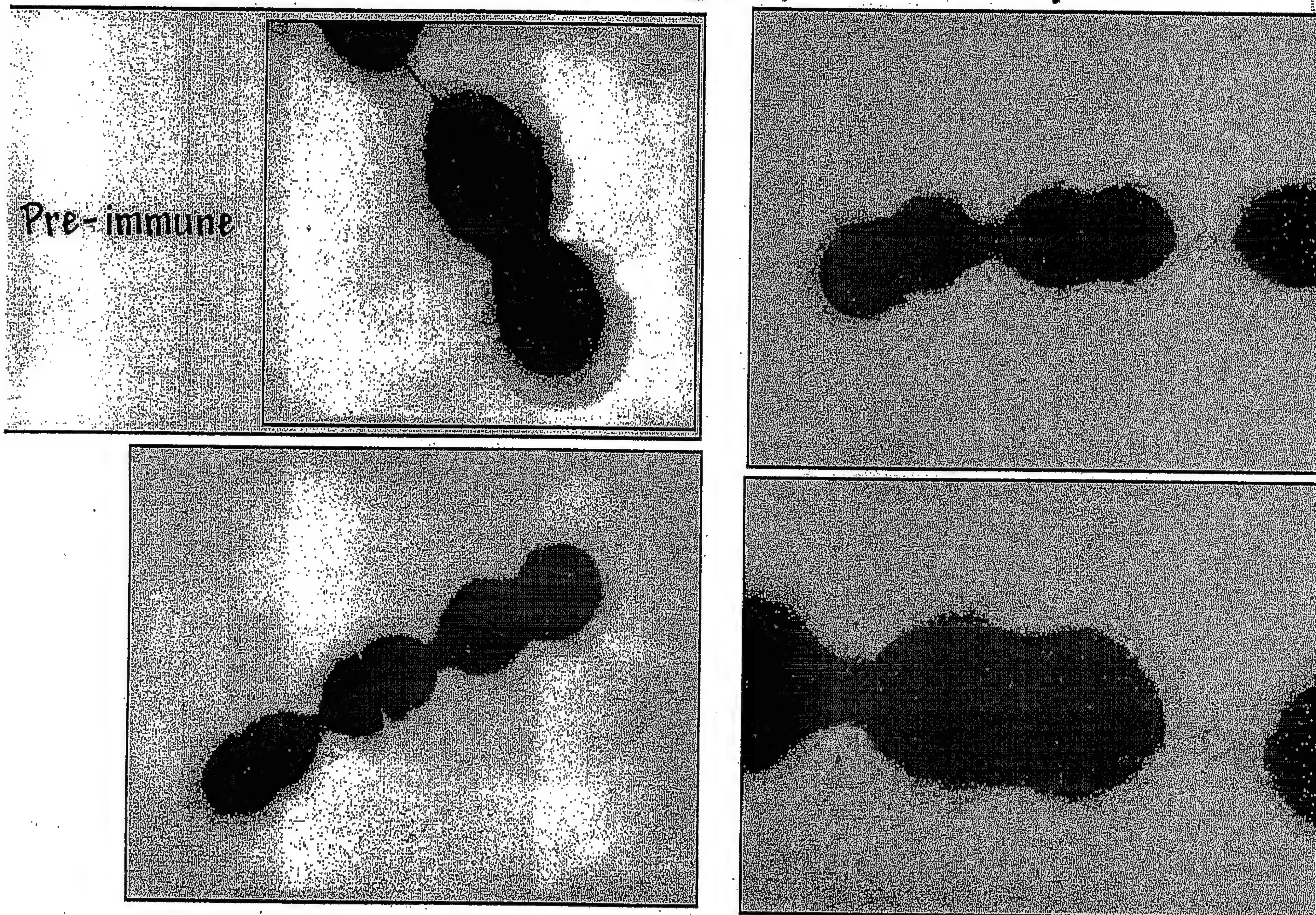


Figure 39 GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)

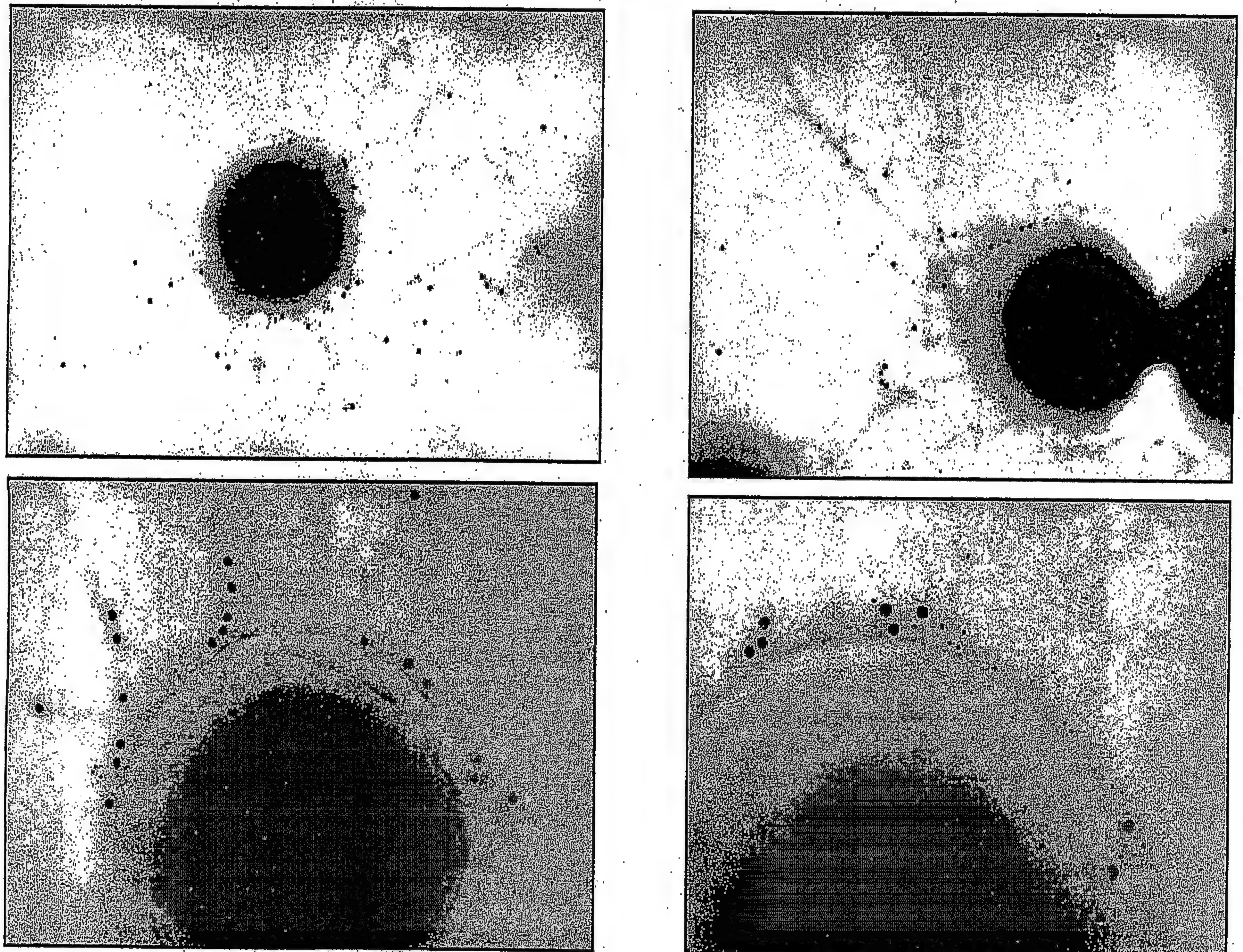


Figure 40

GBS STRAIN COH1 over GBS80

IEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)

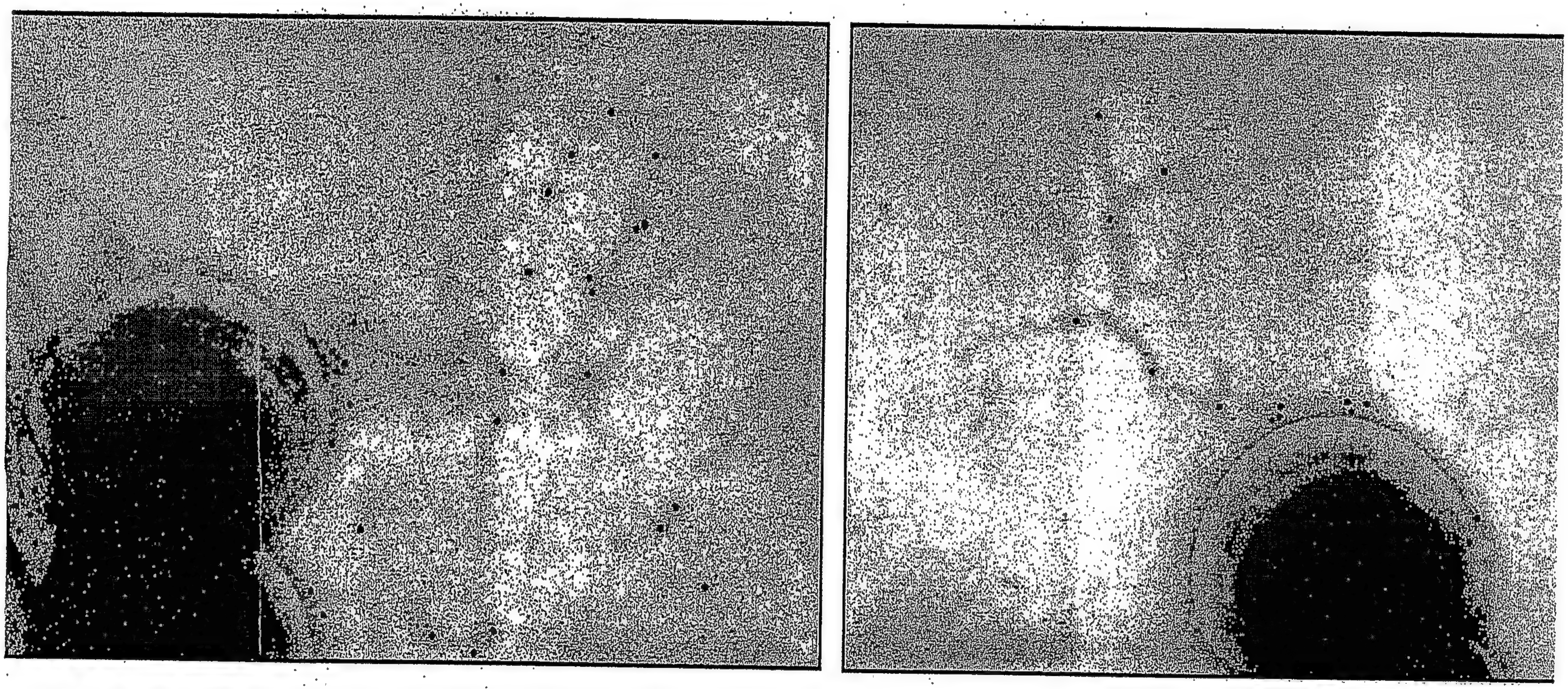


Figure 41: GBS 80 is necessary for polymer formation, GBS104 and sortase SAG0648 are necessary for efficient assembly

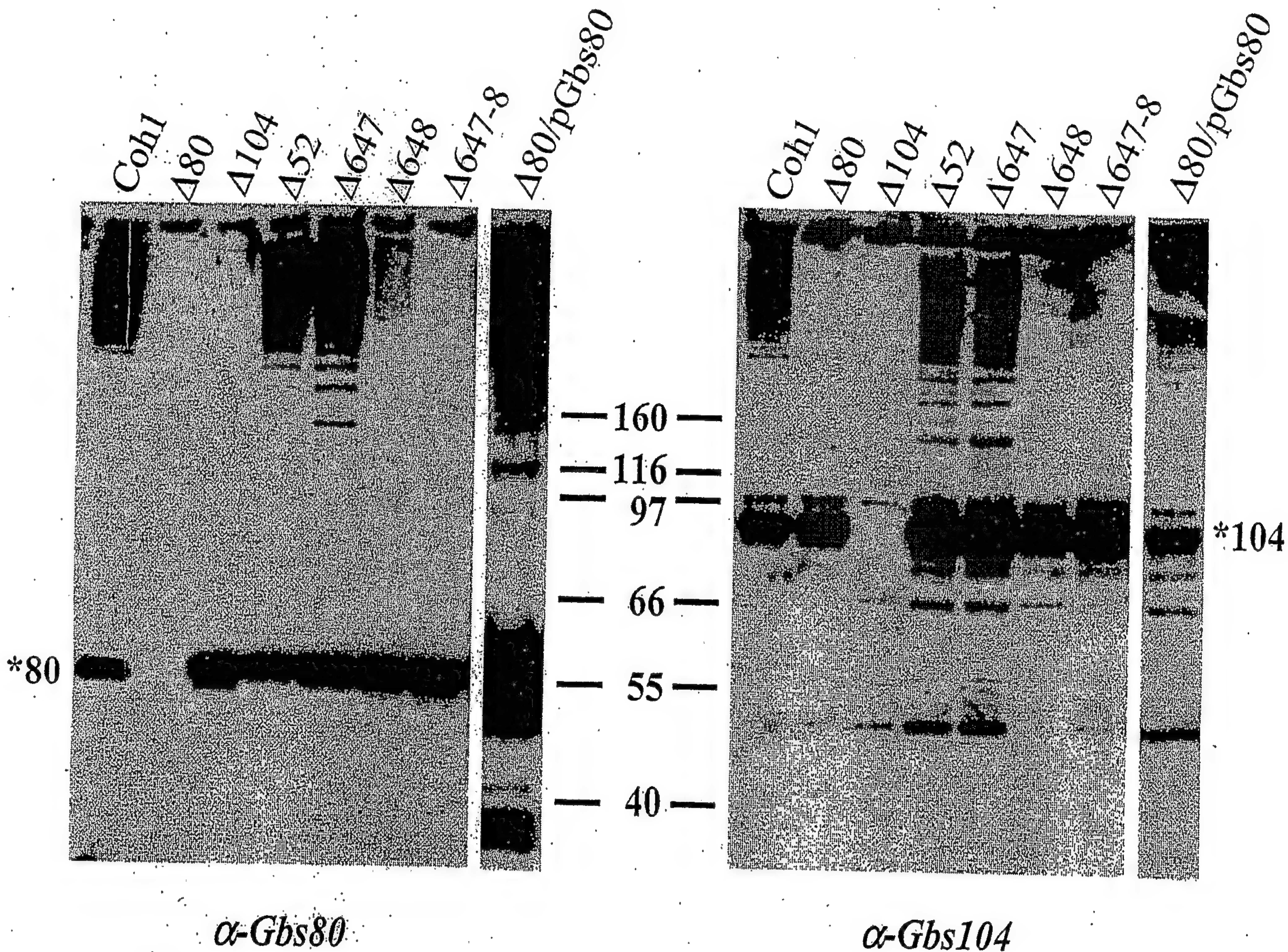


Figure 42: Gbs67 is part of a second pilus; Gbs80 is polymerized in strain 515

(515 lacks sortase 647-8, but has AI-2 sortases)

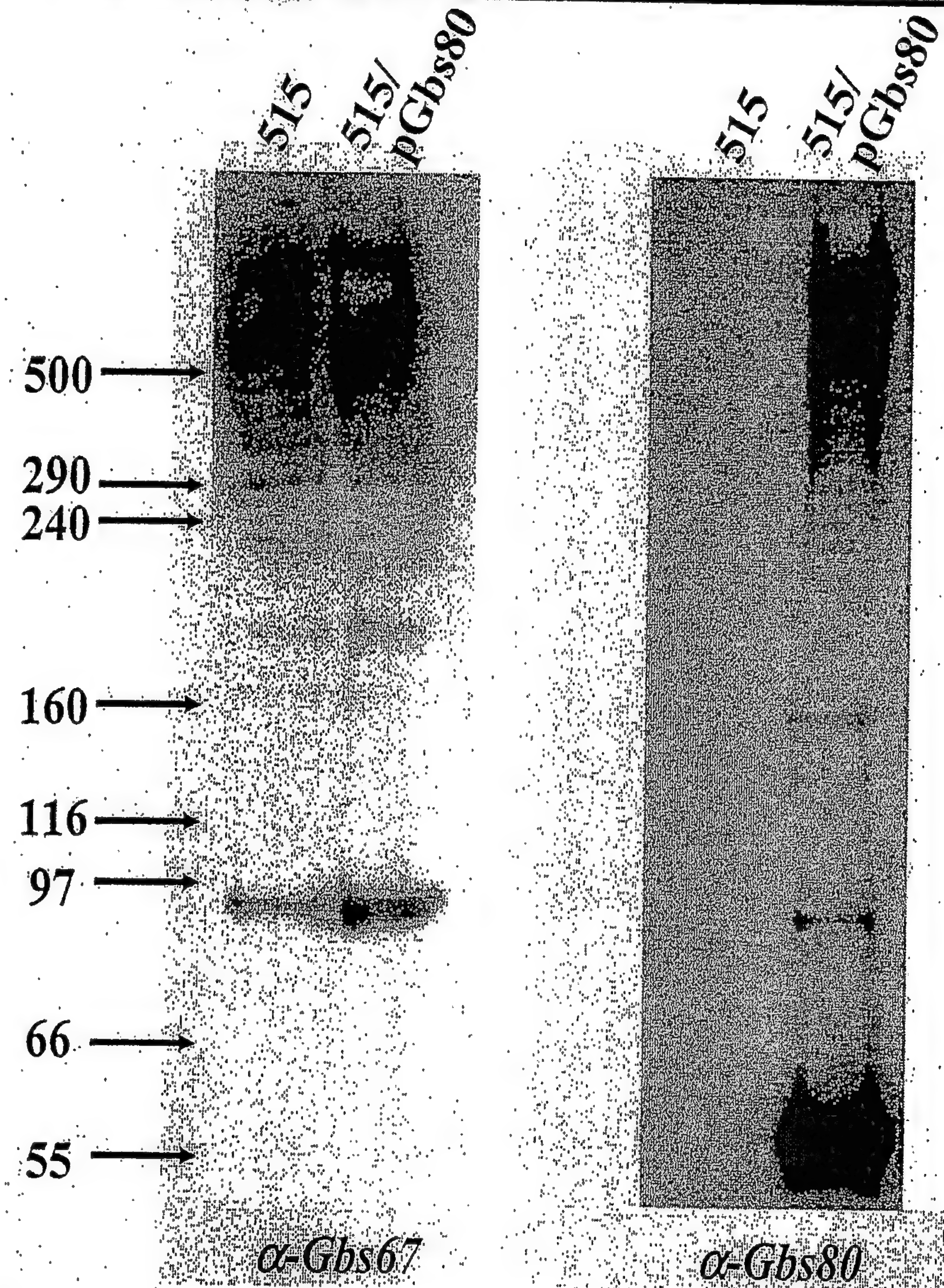
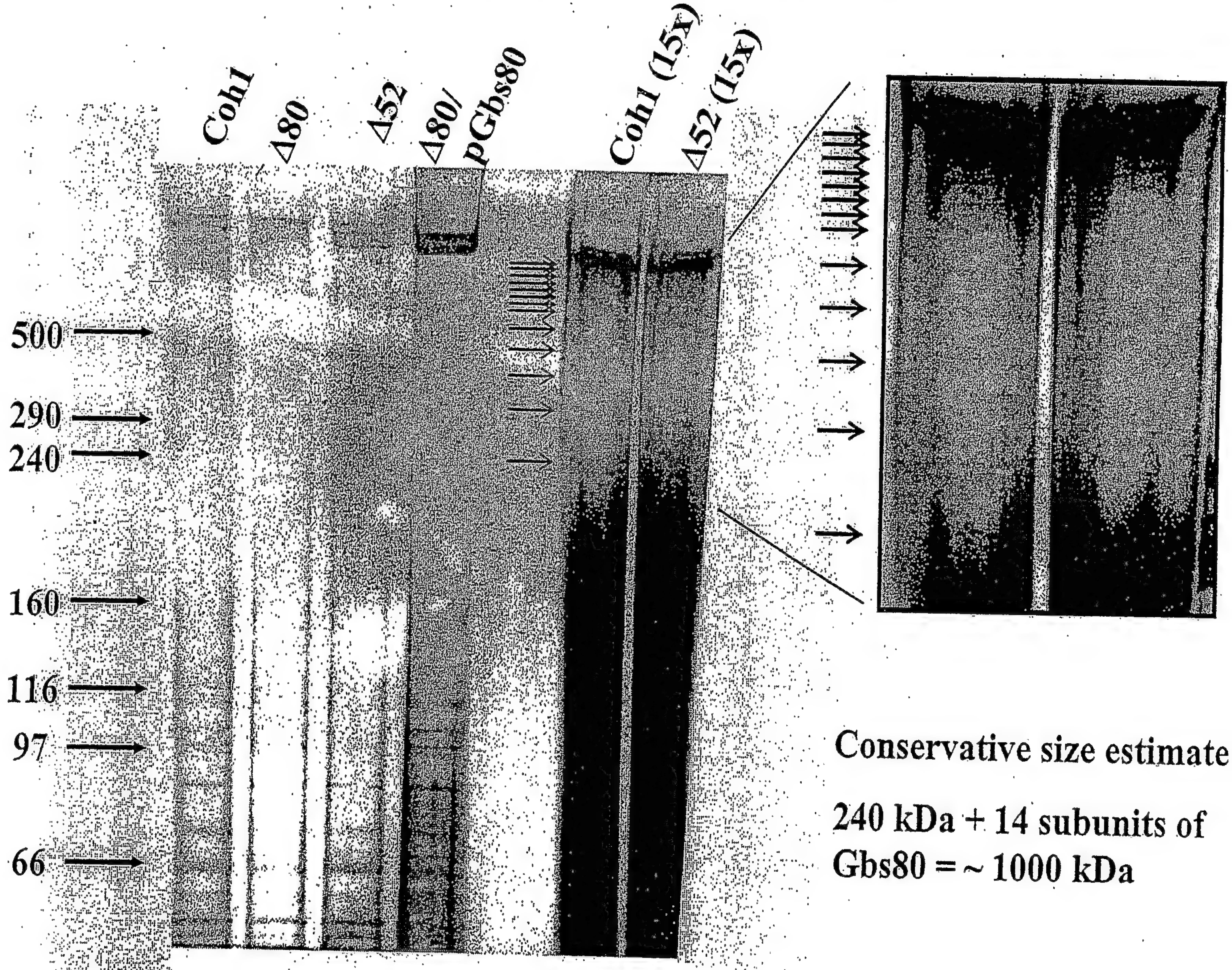


Figure 43: Two macro-molecules are visible in Coh1 at >1000 kDa, one is the Gbs80 pilin



Conservative size estimate:

240 kDa + 14 subunits of Gbs80 = ~ 1000 kDa

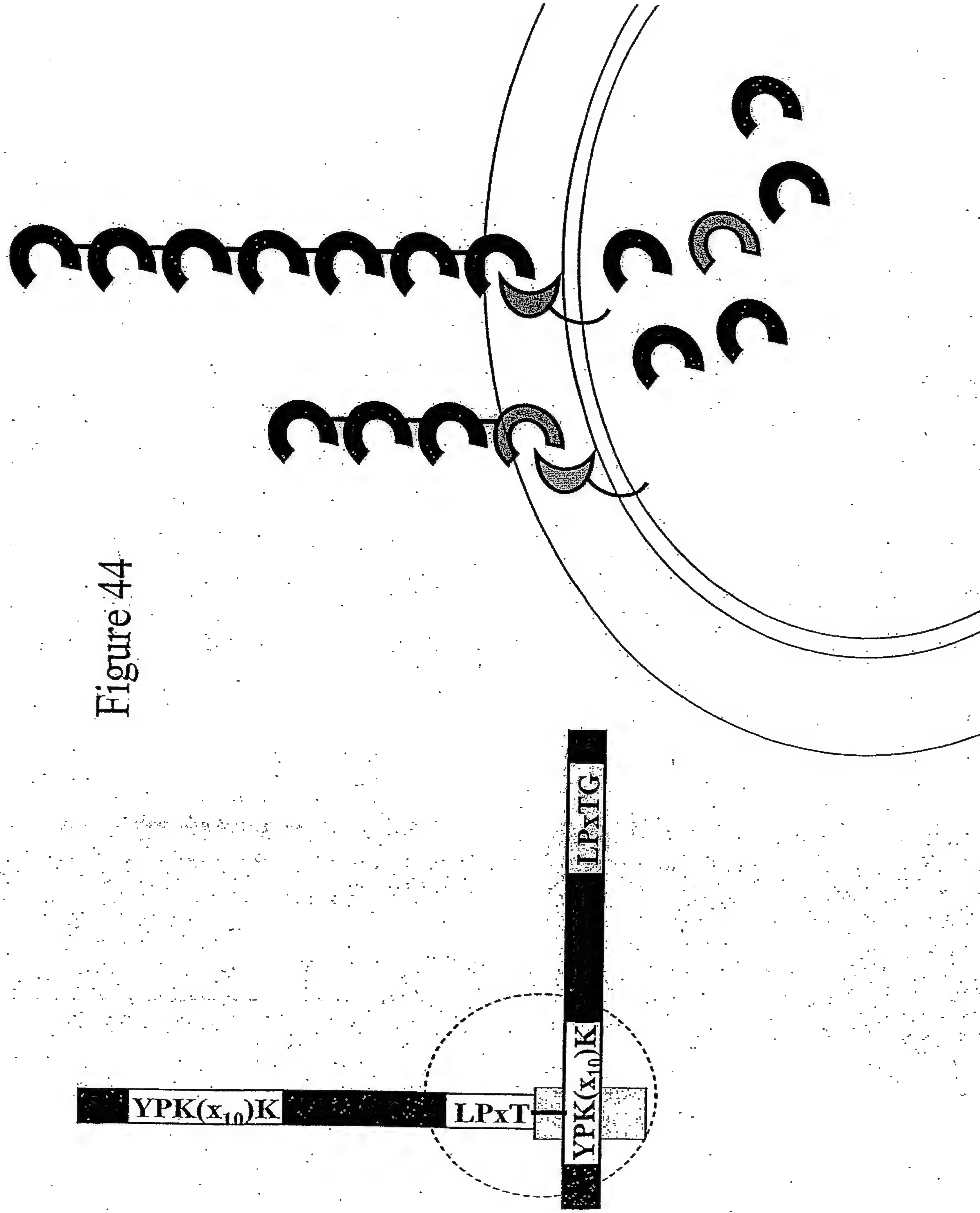
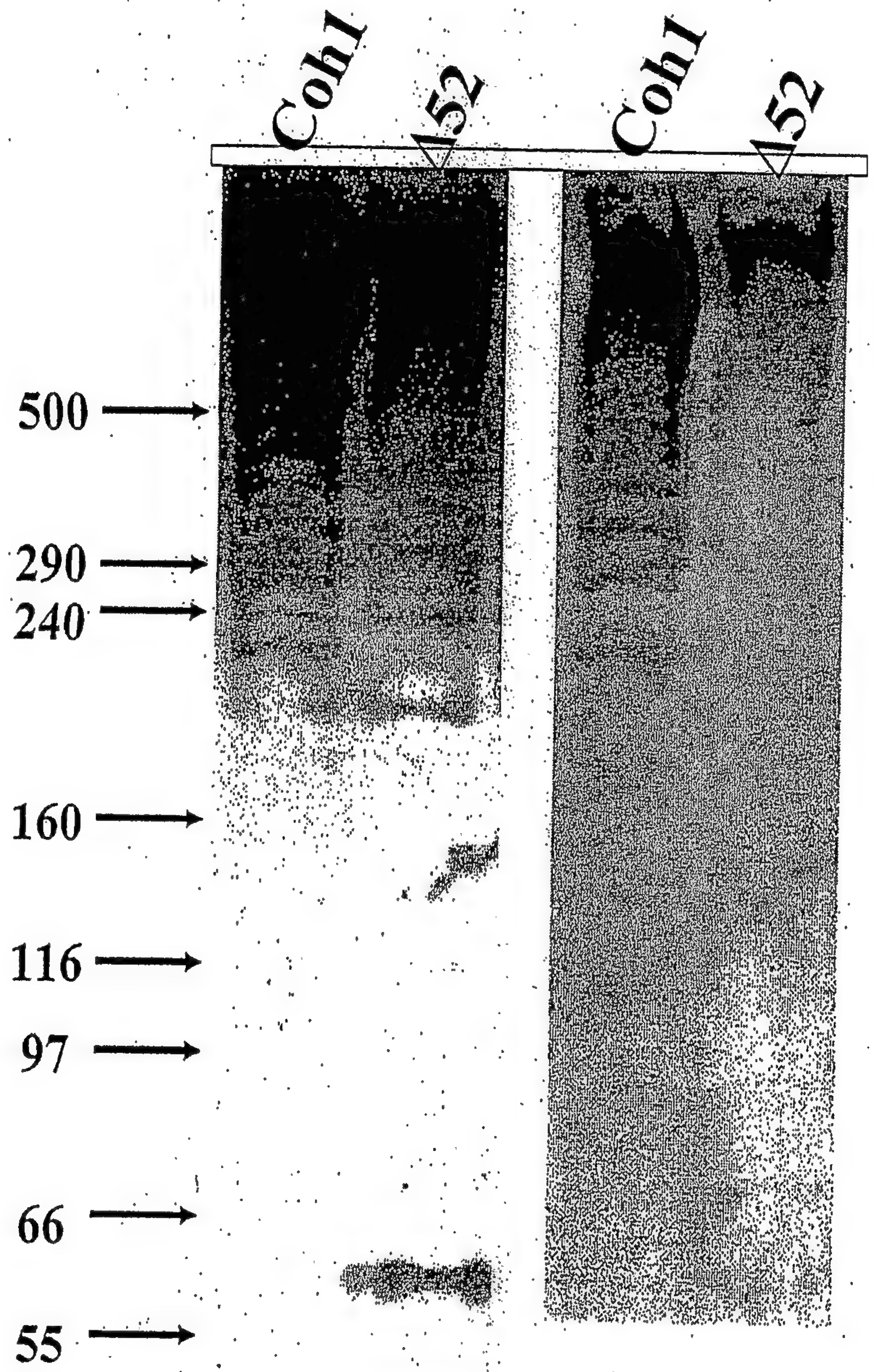


Figure 44

Figure 45: Gbs52 is a minor component of the GBS pilus



3%-8% Tris-acetate gels, includes stack

Left: α-gbs80

Right: blot stripped and reprobed with α-gbs52

Figure 46: The pilus is found in the supernatant of the bacterial culture

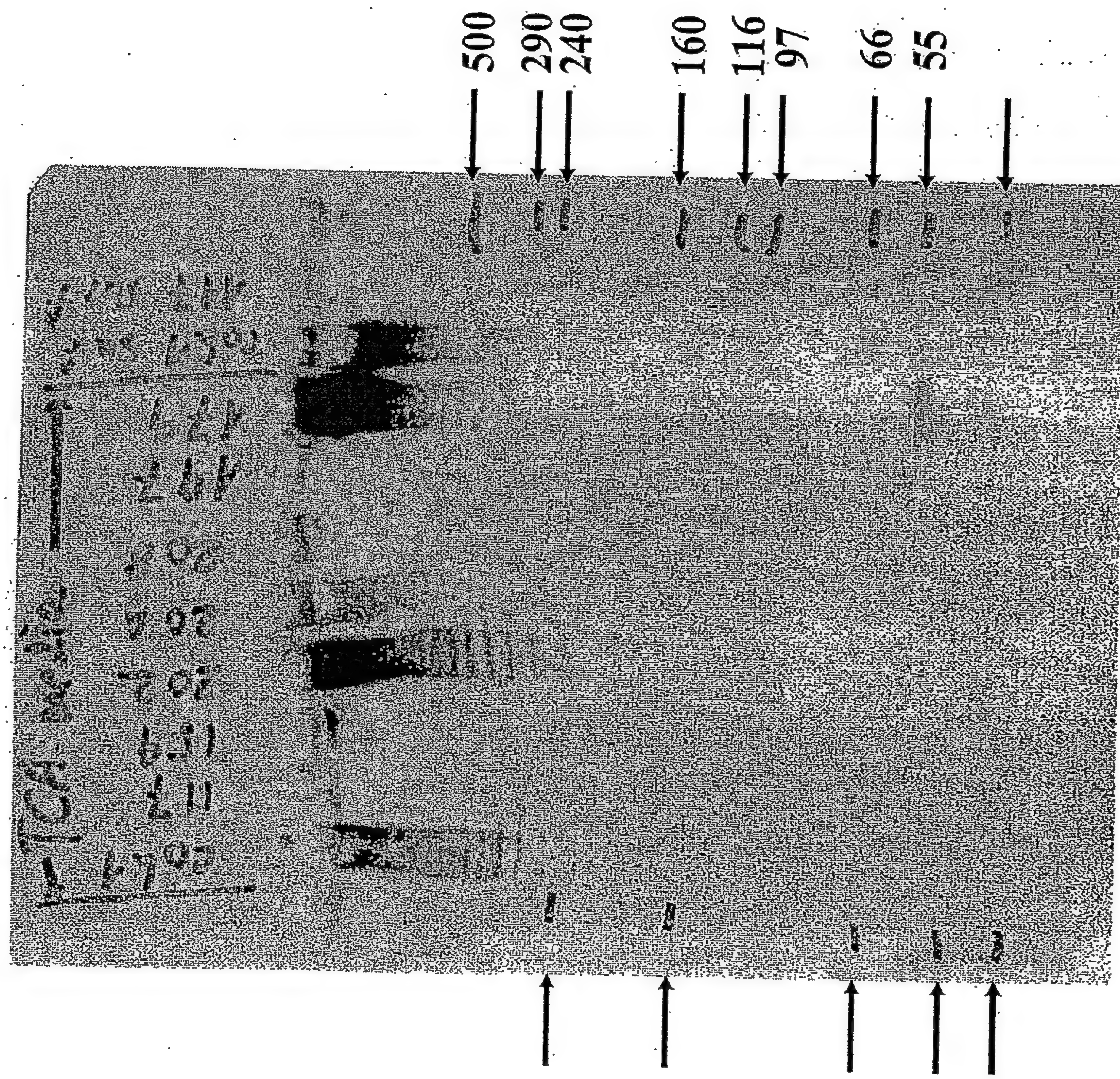
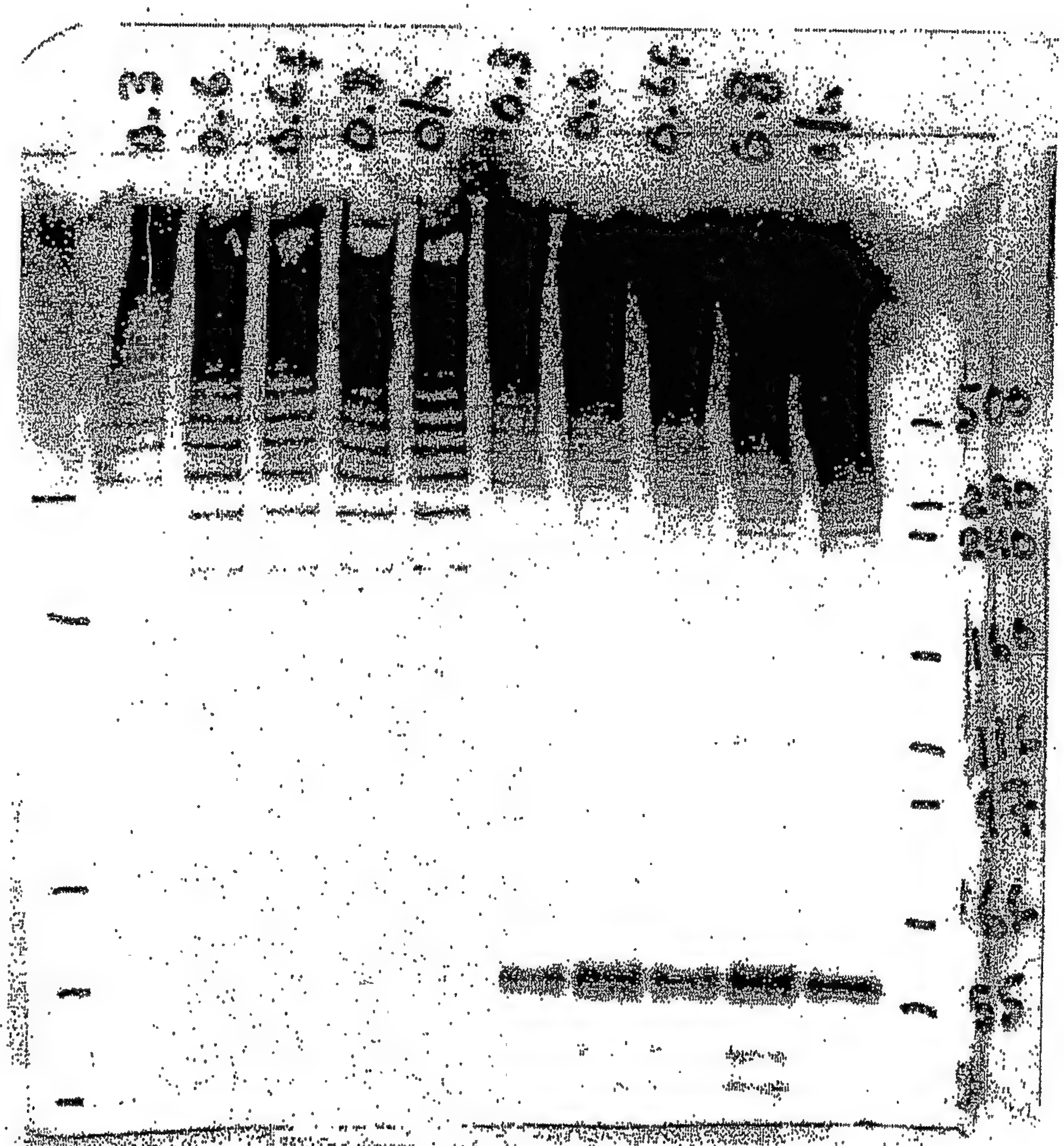


Figure 47: The pilus is found in the supernatant of cultures in all growth phases

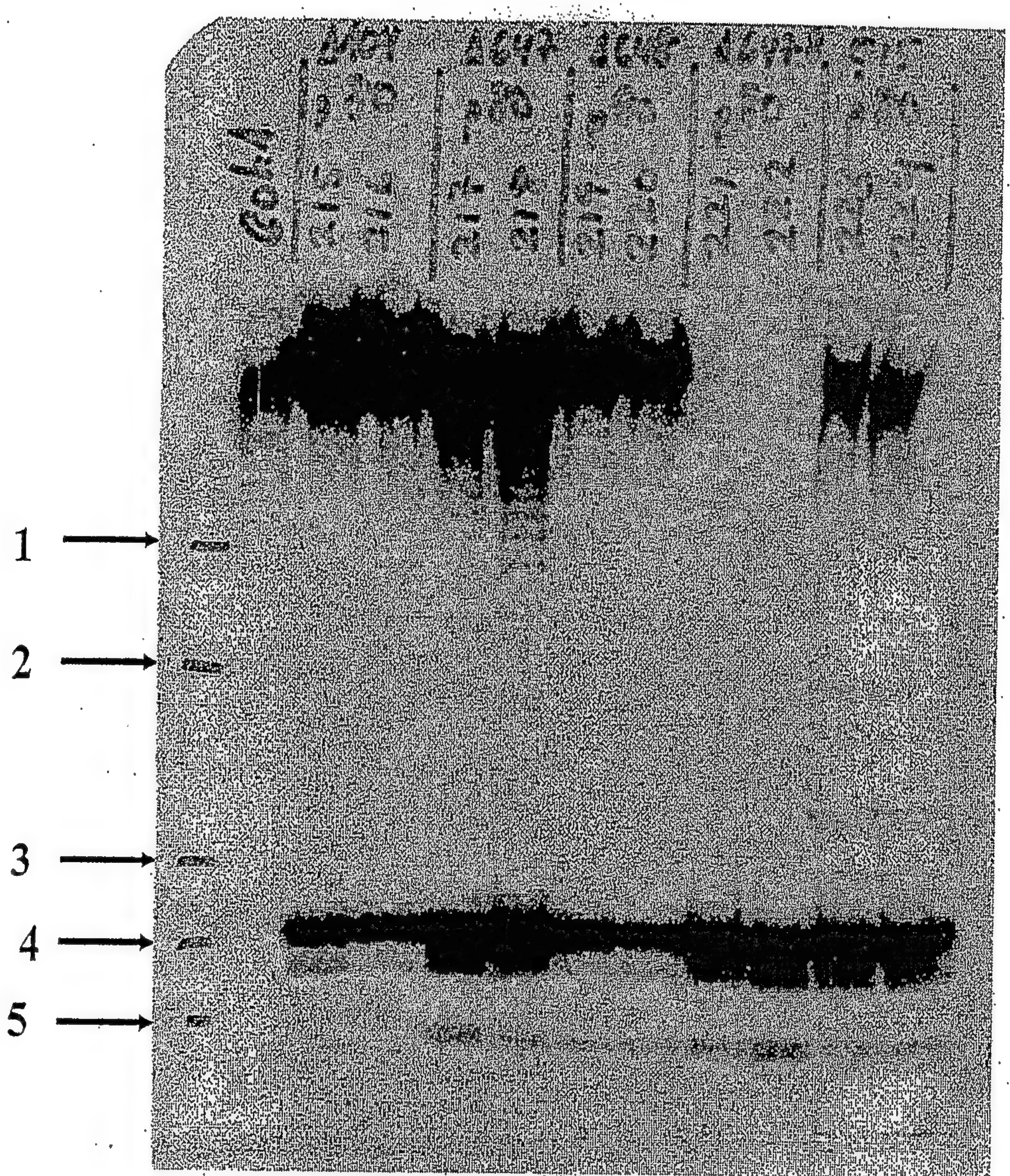


TCA precipitation of 1 ml of THB culture supernatant run on 3-8% SDS-PAGE. OD600 nm are noted above samples, "f" indicates supernatant was filtered (0.2 µM syringe filter).

Left five samples: Coh1.

Right five samples: 179 (ΔGbs80/pGbs80).

Figure 48: In Coh1, only the gbs80 protein and one sortase (sag0647 <u>or</u> sag0648) is required for polymerization



Over expression of gbs80 in various strain backgrounds (two clones each).

Total protein extract preparations.

Only the double sortase mutant does not polymerize gbs80.

Gbs80 is polymerized in the DK515 strain background (lacks adhesin island 1, adhesin island 2 is 2603-like). Presumably, sag1405&sag1406 are responsible for polymerization.

PCT/US05/27239

GBS STRAIN JM9030013
IEM anti-GBS80

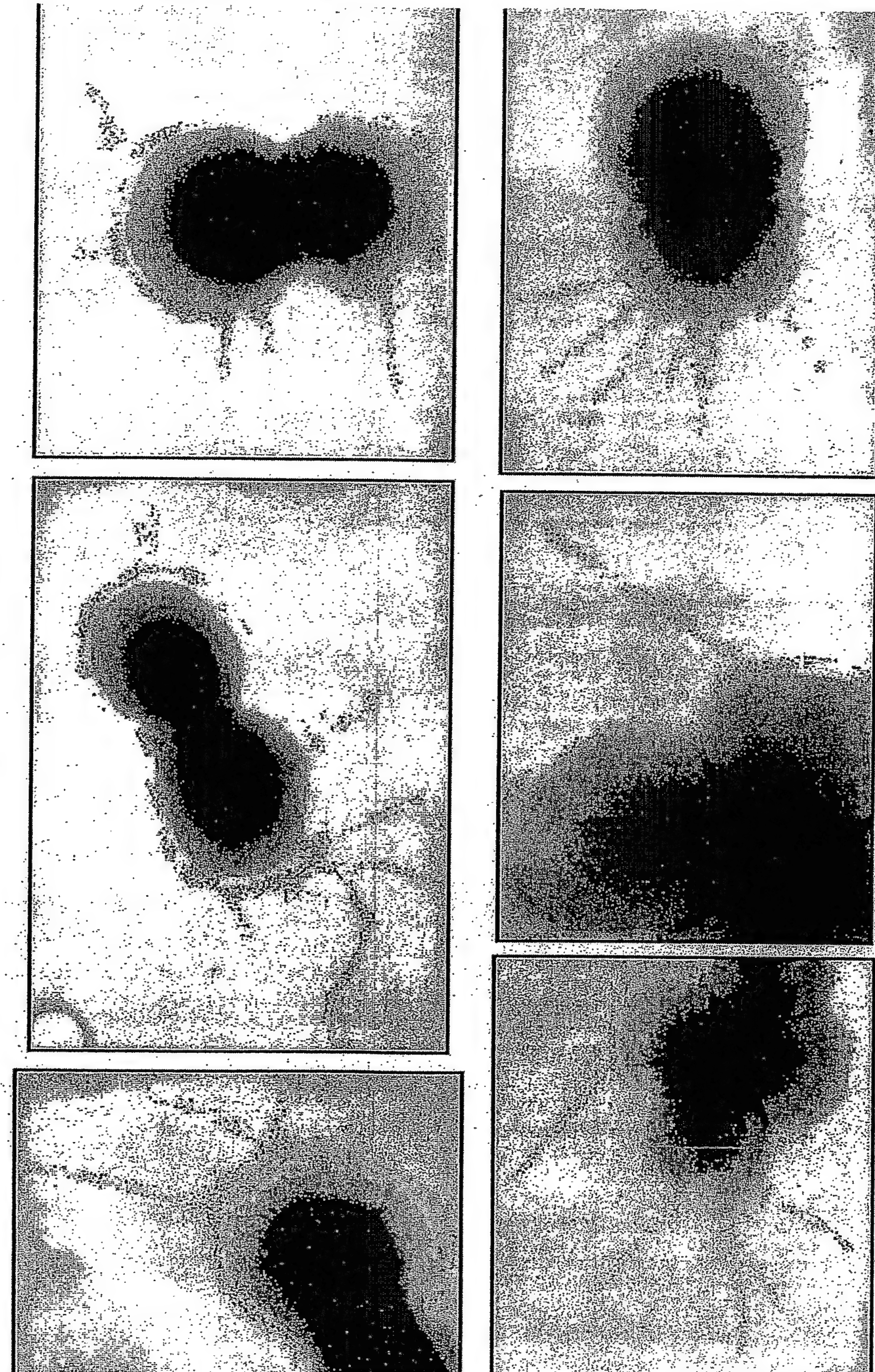


FIGURE 49

PCT/US05/27239

GBS STRAIN JM9030013
IEM anti-GBS104

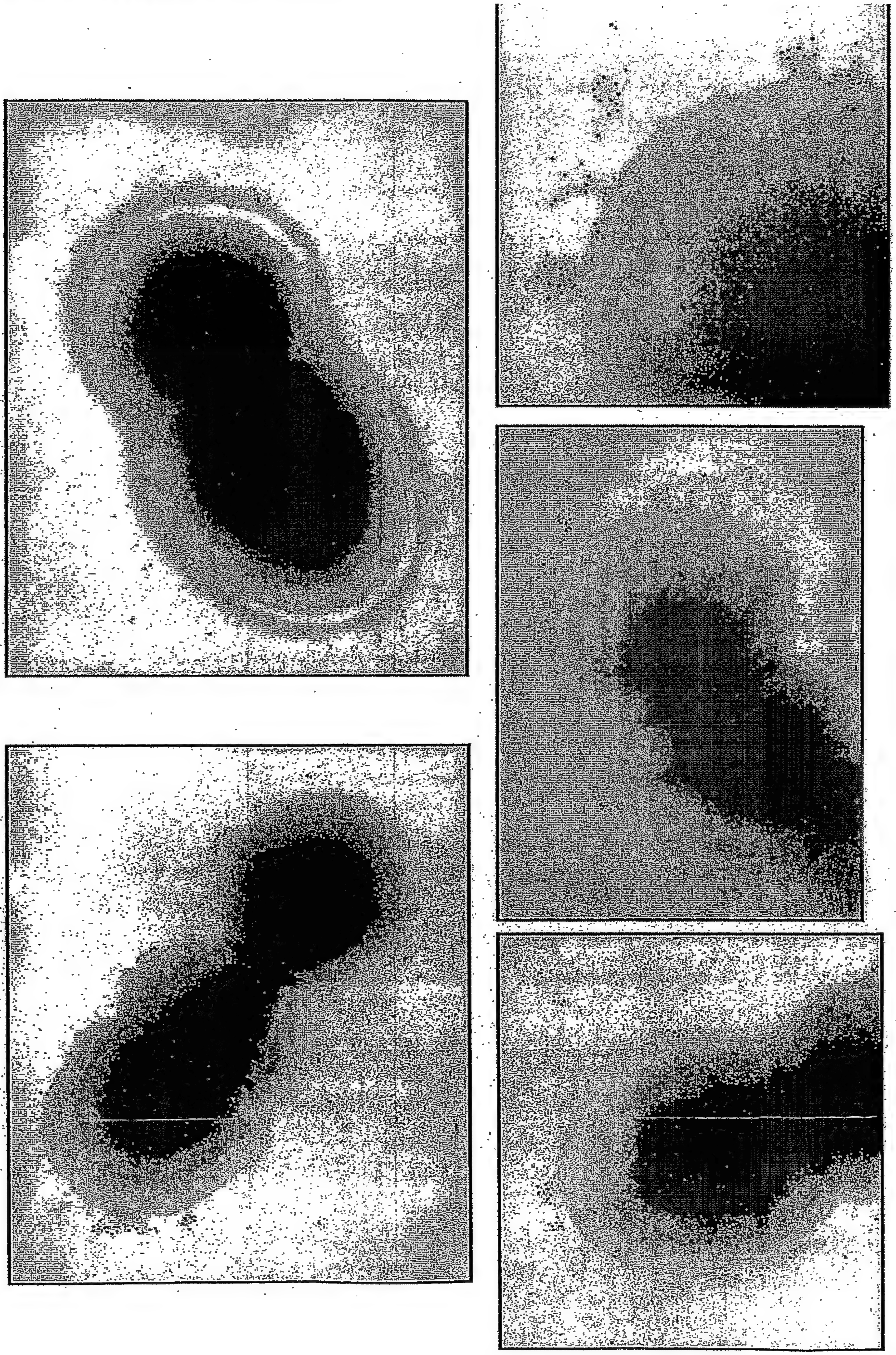
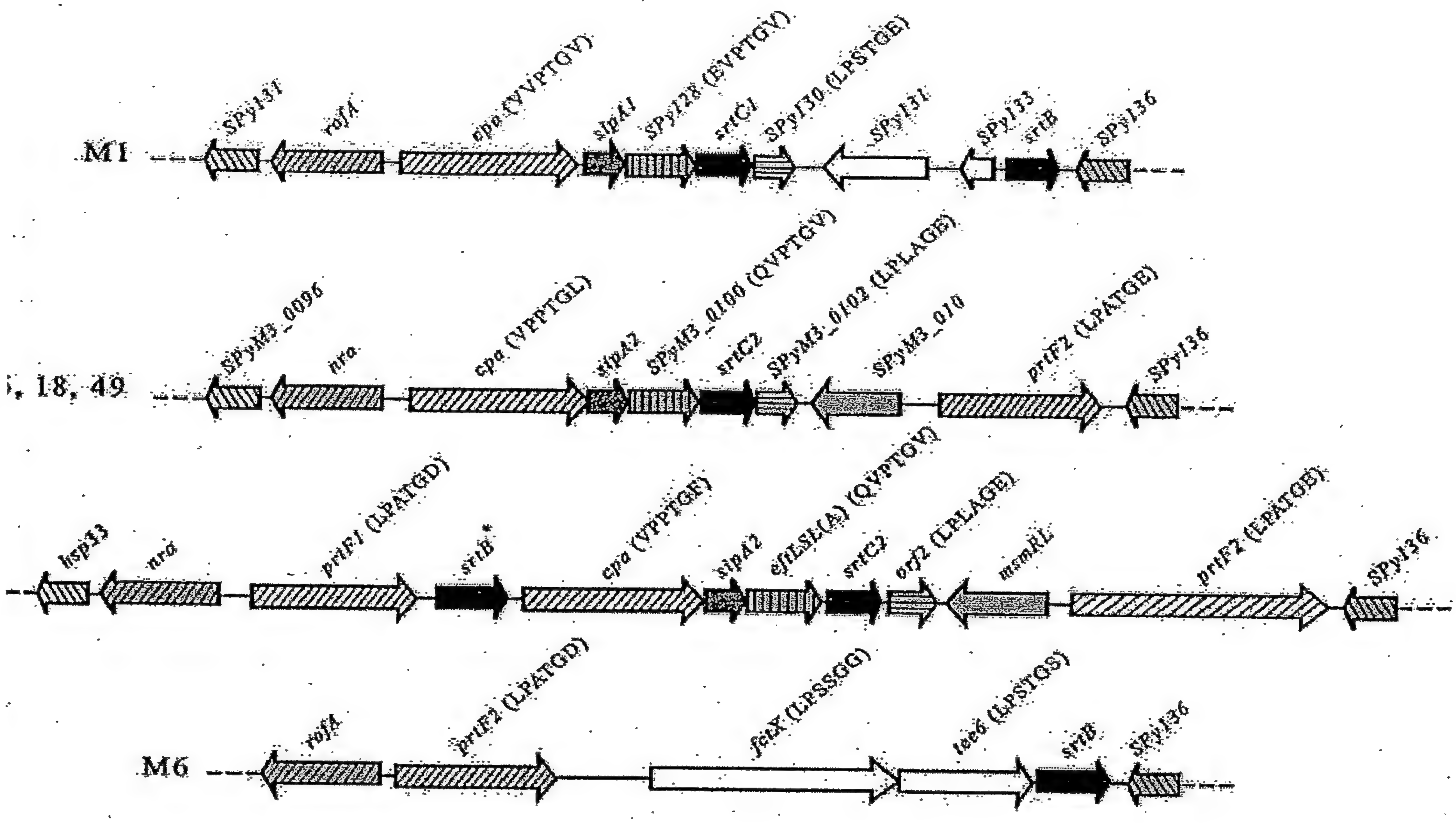


FIGURE 50


M1
SPy131
rofA
cpa (VVPTGV)
sipA1
SPy128 (EVPTGV)
srtC1
SPy130 (LPSTGE)
SPy131
SPy133
srtB
SPy136
, 18, 49
SPyM3_0096
nra
cpa (VPPTGL)
sipA2
SPyM3_0100 (QVPTGV)
srtC2
SPyM3_0102 (LPLAGE)
SPyM3_010
prtF2 (LPATGE)
SPy136
hsp33
nra
prtF1 (LPATGD)
srtB*
cpa (VPPTGE)
sipA2
eftLSL(A) (QVPTGV)
srtC2
orf2 (LPLAGE)
msmRL
prtF2 (LPATGE)
SPy136
M6
rofA
prtF2 (LPATGD)
fctX (LPSSGG)
tee6 (LPSTGS)
srtB
SPy136


M1 1
M3 1
M5 1
M12 1
M18 1
M49 1
M1 80
M3 82
M5 82
M12 82
M18 82
M49 82
M1 161
M3 163
M5 163
M12 163
M18 163
M49 163

FIGURE 51

PCT/US05/27239

```
GI-19224135    1 MNNKKLQKKQDAPR-VSNRKP--------KQLTVTLVGVFLMFLTLVSSMRGAQSIFGEEK
ORF78          1 -----LQKRDKTNYGSANNKR--------RQTTIGLLKVFLTFVALIG-------IVG---
GI-21909634    1 -----MQKRDKTNYGSANNKR--------RQTTIGLLKVFLTFVALIGIVGFSIRAFG---
GI-28810257    1 -----MQKRDKTNYGSANNKR--------RQTTIGLLKVFLTFVALIGIVGFSIRAFG---
GI-19745301    1 -----MQKRDKTNYGSANNKR--------RQTTIGLLKVFLTFVALIGIVGFSIRAFG---
GAS15          1 LRGEKMKKTRFPNKLNTLNTQRVLSKNSKRFTVTLVGVFLMIFALVTSMVGAKTVFG---

GI-19224135   53 RIEEVSVPKIKSPDD--AYPWYGYDSYDSSHPYYERFKVAHDLRVNLNGSKSYQVYCFNI
ORF78         39 ------------------------------------------------------------
GI-21909634   46 -AEEQSVPNKQSSVQ--DYPWYGYDSYSKGYPDYSPLKTYHNLKVNLDGSKEYQAYCFNL
GI-28810257   46 -AEEQSVPNKQSSVQ--DYPWYGYDSYSKGYPDYSPLKTYHNLKVNLDGSKEYQAYCFNL
GI-19745301   46 -AEEQST-----------------------------------------------------
GAS15         58 -LVESSTPNAINPDSSSEYRWYGYESYVRGHPYYKQFRVAHDLRVNLEGSRSYQVYCFNL

GI-19224135  111 NSHYPNRKNAFSKQWFKRVDGTGEVFTNYAQTPKIRGESLNNKLLSIMYNAYPKNANGYM
ORF78         39 ------------------------------------------------------------
GI-21909634  103 TKHFPSKSDSVRSQWYKKLEGTNENFIKLADKPRIEDGQLQQNILRILYNGYPNDRNGIM
GI-28810257  103 TKHFPSKSDSVRSQWYKKLEGTNENFIKLADKPRIEDGQLQQNILRILYNGYPNDRNGIM
GI-19745301   52 ------------------------------------------------------------
GAS15        117 KNAFPLGSDSSVKKWYKKHDGISTKFEDYAMSPRITGDELNQKLRAVMYNGHPQNANGIM

GI-19224135  171 DKIEPLNAILVTQQAVWYYSDSSYGN-IKTLWASELKDGKIDFEQVKLMREAYSKLISDD
ORF78         39 -----------------FSIRAFG------------------------------------
GI-21909634  163 KGIDPLNAILVTQNAIWYYTDSSYISDTSKAFQQEETDLKLDSQQLQLMRNALKRLINPK
GI-28810257  163 KGIDPLNAILVTQNAIWYYTDSSYISDTSKAFQQEETDLKLDSQQLQLMRNALKRLINPK
GI-19745301   52 ------------------------------------------------------------
GAS15        177 EGLEPLNAIRVTQEAVWYYSDNAPISNPDESFKRESESNLVSTSQLSLMRQALKQLIDPN

GI-19224135  230 LEETSKNKLPQGSKLNIFVPQDKS-------VQNLLSAEYVPESPPAPGQSPEPPVQTKKT
ORF78         46 ------------------------------------------------AEEKSTETKKT
GI-21909634  223 EVESLPNQVPANKQLSIFQSSDKT-------FQNLLSAEYVPDTPPKPG--EEPPAKTEKT
GI-28810257  223 EVESLPNQVPANKQLSIFQSSDKT-------FQNLLSAEYVPDTPPKPG--EEPPAKTEKT
GI-19745301   52 -------------------------------------------------------ETKKT
GAS15        237 LATKMPKQVPDDFQLSIFESEDKGDKYNKGYQNLLSGGLVPTKPPTPGDPPMPPNQPQTT

GI-19224135  284 SVIIRKYAEGDYSKLLEGATLRLTGEDILDFQEKVFQSNGTGEKIELSNGTYTLTETSSP
ORF78         57 SVIIRKYAEGDYSKLLEGATLRLTGEDIPDFQEKVFQSNGTGEKIELSNGTYTLTETSSP
GI-21909634  275 SVIIRKYAEGDYSKLLEGATLKLAQIEGSGFQEKIFDSNKSGEKVELPNGTYVLSELKPP
GI-28810257  275 SVIIRKYAEGDYSKLLEGATLKLAQIEGSGFQEKIFDSNKSGEKVELPNGTYVLSELKPP
GI-19745301   57 SVIIRKYAEGDYSKLLEGATLKLAQIEGSGFQEQSFESSTSGQKLQLSDGTYILTETKSP
GAS15        297 SVLIRKYAIGDYSKLLEGATLQLTGDNVNSFQARVFSSNDIGERIELSDGTYTLTELNSP

GI-19224135  344 DGYKIAEPIKFRVVNKKVFIVQKDGSQVENPNKEVAEPYSVEAYSDMQDSNYINPETFTP
ORF78        117 DGYKITEPIKFRVVNKKVFIVQKDGSQVENPNKELGSPYTIDAYNDFDEEGLLSTQN---
GI-21909634  335 QGYGVATPITFKVAAEKVLIKNKEGQFVENQNKEIAEPYSVTAFNDFEEIGYLS--DFNN
GI-28810257  335 QGYGVATPITFKVAAEKVLIKNKEGQFVENQNKEIAEPYSVTAFNDFEEIGYLS--DFNN
GI-19745301  117 QGYEIAEPITFKVTAGKVFINGKDGQFVENQNKEVAEPYSVIAYNDFDDSGFINPKTFTP
GAS15        357 AGYSIAEPITFKVEAGKVYT-IIDGKQIENPNKEIVEPYSVEAYNDFEEFSVLTTQN---

GI-19224135  404 YGKFYYAKNKDKSSQVVYCFNADLHSPPESEDGGGTIDPDISTMKEVKYTHIAGSDLFKY
ORF78        174 YAKFYYGKNYDGSSQIVYCFNANLKSPPDSEDHGATINPDFTTG-DIRYSHIAGSDLIKY
GI-21909634  393 YGKFYYAKNTNGTNQVVYCFNADLHSPPDSYDHGANIDPDVSESKEIKYTHVSGYDLYKY
GI-28810257  393 YGKFYYAKNTNGTNQVVYCFNADLHSPPDSYDHGANIDPDVSESKEIKYTHVSGYDLYKY
GI-19745301  177 YGKFYYAKNANGTSQVVYCFNVDLHSPPDSLDKGETLDPDFNEGKEIKYTHILGADLFSY
GAS15        413 YAKFYYAKNKNGSSQVVYCFNADLKSPPDSEDGGKTMTPDFTTG-EVKYTHIAGRDLFKY
```

FIGURE 52

```
GI-19224135  464 ALRPRDTNPEDFLKHIKKVIEKGYNKKGD--SYNGLTETQFRAATQLAIYYFTDSTDLKT
ORF78        233 ANTARDEDPQLFLKHVNKVIENGYHKKGQAIPYNGLTEAQFRAATQLAIYYFTDSVDL--
GI-21909634  453 AATPRDKDADFFLKHIKKILDKGYKKKGD--TYKTLTEAQFRAATQLAIYYYTDSADLTT
GI-28810257  453 AATPRDKDADFFLKHIKKILDKGYKKKGD--TYKTLTEAQFRAATQLAIYYYTDSADLTT
GI-19745301  237 ANNPRASTNDELLSQVIKIVLEKGYPDDST--TYANLTSVEFRAATQLAIYYFTDSVDLDN
GAS15        472 TVKPRDTDPDTFLKHIKKVIEKGYREKGQAIEYSGLTETQLRAATQLAIYYFTDSAELD-

GI-19224135  522 LKTYNNGKGYHGFESMDEKTLAVTKELINYAQD-NSAPQLTNLDFFVPNNSKYQSLIGTE
ORF78        291 --TKDRLKDEHGFGDMNDQTLGVAKKIVEYALS-DEDSKLTNLDFFVPNNSKYQSLIGTE
GI-21909634  511 LKTYNDNKGYHGFDKLDDATLAVVHELITYAED-VTLPMTQNLDFFVPNSSRYQALIGTQ
GI-28810257  511 LKTYNDNKGYHGFDKLDDATLAVVHELITYAED-VTLPMTQNLDFFVPNSSRYQALIGTQ
GI-19745301  295 LADY-------HGFGALTTEALNATKEIVAYAEDRANLPNISNLDFYVPNSNKYQSLIGTQ
GAS15        531 ---KDKLKDYHGFGDMNDSTLAVAKILVEYAQD-SNPPQLTDLDFFIPNNNKYQSLIGTQ

GI-19224135  581 YHPDDLVDVIRMEDKKQEVIPVTHSLTVKKTVVGELGDKTKGFQFELELKDKTGQPIVNT
ORF78        348 YHPDDLVDVIRMEDKKQEVIPVTHSLTVQKTVVGELGDKTKGFQFELELKDKTGQPIVNT
GI-21909634  570 YHPNELIDVISMEDKQAPIIPITHKLTISKTVTGTIADKKKEFNFEIHLKSSDGQAISGT
GI-28810257  570 YHPNELIDVISMEDKQAPIIPITHKLTISKTVTGTIADKKKEFNFEIHLKSSDGQAISGT
GI-19745301  349 YHPESLVDIIRMEDKQAPIIPITHKLTISKTVTGTIADKKKEFNFEIHLKSSDGQAISGT
GAS15        587 WHPEDLVDIIRMEDKK-EVIPVTHNLTIRKTVTGLAGDRTKDFHFEIELKNNKQELLSQT

GI-19224135  641 LKTNNQDLVAKDGKYSFNLKHGDTIRIEGLPTGYSYTLKETEAKDYIVTVDNKVSQEAQS
ORF78        408 LKTNNQDLVAKDGKYSFNLKHGDTIRIEGLPTGYSYTLKETEAKDYIVTVDNKVSQEAQS
GI-21909634  630 YPTNSGELTVTDGKATFTLKDGESLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GI-28810257  630 YPTNSGELTVTDGKATFTLKDGESLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GI-19745301  409 YPTNSGELTVTDGKATFTLKDGESLIVEGLPSGYSYEITETGASDYEVSVNGKNAPDGKA
GAS15        646 VKTDKTNLEFKDGKATINLKHGESLTLQGLPEGYSYLVKETDSEGYKVKVNSQEVANATV

GI-19224135  701 ASENVTADKEVTFENRKDLVPPTGFITDGGTYLWLLLLVPFGLLVWFFGRKGLKND-
ORF78        468 ASENVTADKEVTFENRKDLVPPTGLTTDGAIYLWLLLLVPFGLLVWLFGRKGLKND-
GI-21909634  690 TKASVKEDETVAFENRKDLVPPTGLTTDGAIYLWLLLLVPFGLLVWLFGRKGIKK--
GI-28810257  690 TKASVKEDETVAFENRKDLVPPTGLTTDGAIYLWLLLLVPFGLLVWLFGRKGIKK--
GI-19745301  469 TKASVKEDETITFENRKDLVPPTGLTTDGAIYLWLLLLVLLGLWVWLTGRKGLKND-
GAS15        706 SKTGITSDETLAFENNKEPVVPTGVDQKINGYLALIVTAGISLGIIGIHTIRIRKHD
```

FIGURE 52A

```
GI-1922...      1   ..SSYMFARGEKMNNKMFLNKEAGFLVHTKRKRKFAVTLVGVFFLLLACAGAIGFGQVAY
GI-50913503     1   MVSSYMFVRGEKMNNKIFLNKEASFLAHTKRKRRFAVTLVGVFFMLLACAGAIGFGQVAY

GI-19224134    61   AADEKTVPNFKSPDPDYPWYGYDSY------RGIFARYHNLKVNLKGSKEYQAYCFNETK
GI-50913503    61   AADEKTVPSHSSPNPEFPWYGYDAYGKEYPGYNIWTRYHDLRVNLNGSRSYQVYCFNIQS

GI-19224134   115   YEPRPTYSTTNNEYKKIDGSGSAFKSYAANPRVLDENLDKLEKNILNVIYNGYKSNANGE
GI-50913503   121   NYPSQKNSFIKNWEKKIEGNGKSFVDYAHTTKLGKE---ELEQRELSLEYNEYPNDANGY

GI-19224134   175   MNGIEDLNAILVTQNAIIYYSDSAPLNDVNKMWEREVRNGEISESQVTLMREALKKLIDP
GI-50913503   178   MKGLEHLNAITVTQYAVIHYSDNS-QYQFETLWESEANEGKISRSQVTLMREALKKLIDP

GI-19224134   235   NLEATAANKIPSGYRLNIFKSENEDYQNLLSAEYVPDDPPKPGDTSEHNPKTPELDGTPI
GI-50913503   237   NLEATAVNKIPSGYRLNIFESENEAYQNLLSAEYVPDDPPKPGETSEHNPKTPELDGTPI

GI-19224134   295   PEDPKRPDESSEPALPPLMPELDGEEVPEVPSESLEPALPPLMPELDGEEVPEVPSESLE
GI-50913503   297   PEDPKHPDDNLEPTLPPVM-----------------------------------------

GI-19224134   355   PALPPLMPELDGEEVPEVPSESLEPALPPLMPELDGEEVPEVPSESLEPALPPLMPELDG
GI-50913503   316   -----------------------------------LDGEEVPEVPSESLEPALPPLMPELDG

GI-19224134   415   EEVPEKPSVDLPIEVPRYEFNNKDQSPLAGESGETEYITEVYGNQQNPVDIDKKLPNETG
GI-50913503   343   QEVPEKPSIDLPIEVPRYEFNNKDQSPLAGESGETEYITEVYGNQQNPVDIDKKLPNETG

GI-19224134   475   FSGNMVETEDTKEPEVLMGGQSESVEFTKDTQTGMSGQTTPQVETEDTKEPEVLMGGQSE
GI-50913503   403   FSGNMVETEDTKEPEVLMGGQSESVEFTKDTQTGMSGQTTPQVETEDTKEPEVLMGGQSE

GI-19224134   535   SVEFTKDTQTGMSGQTTPQVETEDTKEPGVLMGGQSESVEFTKDTQTGMSGQTTPQVETE
GI-50913503   463   SVEFTKDTQTGMSGQTTPQIETEDTKEPEVLMGGQSESVEFTKDTQTGMSGQTTPQIETE

GI-19224134   595   DTKEPGVLMGGQSESVEFTKDTQTGMSGFSETVTIVEDTRPKLVFHFDNNEPKVEENREK
GI-50913503   523   DTKEPEVLMGGQSESVEFTKDTQTGMSGFSETATVVEDTRPKLVFHFDNNEPKVEENREK

GI-19224134   655   PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNKQNNKV-
GI-50913503   583   PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNKQSNKKV
```

FIGURE 53

```
GI-19745307    1  MTQKNSY  FLLSLTGFILGLLLVFIGLSGVSVGHAETRNGANKQG
ORF84             NSY  SFLLSLTGFILGLLLVFIGLSGVSVGHAETRNGANKQGAF
GI-28810263    1  MTQKNSYKLSFLLSLTGFILGLLLVFIGLSGVSVGHAETRNGANKQGAFEIKKN------
GI-21909640    1
GI-19224141    1  MTQKNSYKLSFLLSLTGFILGLLLVFIGLSGVSVGHAETRNGANKQGSFEIKKVDQNNKP

GI-19745307   55  ------------------------------------------------KSQEEYNYE
ORF84         55  ------------------------------------------------KSQEEYNYE
GI-28810263   55  ------------------------------------------------KSQEEYNYE
GI-21909640    1  ------------------------------------------------KSQEEYNYE
GI-19224141   61  LPGATFSLTSKDGKGTSVQTFTSNDKGIVDAQNLQPGTYTLKEETAPDGYDKTSRTWTVT

GI-19745307   64  VYDN-------------------------------------------------RNI
ORF84         64  VYDN-------------------------------------------------RNI
GI-28810263   64  VYDN-------------------------------------------------RNI
GI-21909640    1  -----------------------------------------------------RNI
GI-19224141  121  VYENGYTKLVENPYNGEIISKAGSKDVSSSLQLENPKMSVVSKYGKTEVSSGAADFYRNH

GI-19745307   71  LQDGEHKLEIKRVDGTGKTYQG--FCFQLTKNFP---TAQGVSKKLYKKLSS--------
ORF84         71  LQDGEHKLEIKRVDGTGKTYQG--FCFQLTKNFP---TAQGVSKKLYKKLSS--------
GI-28810263   71  LQDGEHKLEIKRVDGTGKTYQG--FCFQLTKNFP---TAQGVSKKLYKKLSS--------
GI-21909640    1  -------------------------------------------------MSS--------
GI-19224141  181  AAYFKMSFEEKQKDKSETINPGDTEVLQLDRRLNPKGISQDEPKIIYDSANSPLAIGKYH

GI-19745307  118  -----------------------------SDEETLK------------------------
ORF84        118  -----------------------------SDEETLK------------------------
GI-28810263  118  -----------------------------SDEETLK------------------------
GI-21909640    4  -----------------------------SDEETLK------------------------
GI-19224141  241  AENHQLIYTFTDYIAGLDKVQLSAELSLFLENKEVLENTSISNFKSTIGGQEITYKGTVN

GI-19745307  125  -------QYASKYTSNRRGDTSG-------------------------------------
ORF84        125  -------QYASKYTSNRRGDTSG-------------------------------------
GI-28810263  125  -------QYASKYTSNRRGDTSG-------------------------------------
GI-21909640   11  -------QYASKYTSNRRGDTSG-------------------------------------
GI-19224141  301  VLYGNESTKESNYITNGLSNVGGSIESYNTETGEFVWYVYVNPNRTNIPYATMNLWGFGR

GI-19745307  141  -------------------------------------------NLKKQIAKVLTEGYPT
ORF84        141  -------------------------------------------NLKKQIAKVLTEGYPT
GI-28810263  141  -------------------------------------------NLKKQIAKVLTEGYPT
GI-21909640   27  -------------------------------------------NLKKQIAKVLTEGYPT
GI-19224141  361  ARSNTSDLENDANTSSAELGEIQVYEVPEGEKLPSSYGVDVTKLTLRTDITAGLGNGFQM

GI-19745307  157  NKS---DWLN----------------------------GLTENEKIEVTQDAIWYF
ORF84        157  NKS---DWLN----------------------------GLTENEKIEVTQDAIWYF
GI-28810263  157  NKS---DWLN----------------------------GLTENEKIEVTQDAIWYF
GI-21909640   43  NKS---DWLN----------------------------GLTENEKIEVTQDAIWYF
GI-19224141  421  TKRQRIDEGNNIQNKAFIIKVTGKTDQSGKPLVVQSNLASFRGASEYAAFTPVGGNVYFQ

GI-19745307  182  TETTVPADR-------SYTNRNVNSQKMKEVYQKLIDTTDID---KYEDVQFDLFVPQDTN
ORF84        182  TETTVPADR-------SYTNRNVNSQKMKEVYQKLIDTTDID---KYEDVQFDLFVPQDTN
GI-28810263  182  TETTVPADR-------SYTNRNVNSQKMKEVYQKLIDTTDID---KYEDVQFDLFVPQDTN
GI-21909640   68  TETTVPADR-------SYTNRNVNSQKMKEVYQKLIDTTDID---KYEDVQFDLFVPQDTN
GI-19224141  481  NEIALSPSKGSGSGKSEFTKPSITVANLKRVAQLRFKKMSTDNVPLPEAAFELRSSNGNS

GI-19745307  233  --LQAVISVEPVIESLPWTS-----------------LKPIAQKDITAKK----------
ORF84        233  --LQAVISVEPVIESLPWTS-----------------LKPIAQKDITAKK----------
GI-28810263  233  --LQAVISVEPVIESLPWTS-----------------LKPIAQKDITAKK----------
GI-21909640  119  --LQAVISVEPVIESLPWTS-----------------LKPIAQKDITAKK----------
GI-19224141  541  QKLEASSNTQGEVHFKDLTSGTYDLYETKAPKGYQQVTEKEATVTVDTTKPAEEMVTWGS
```

FIGURE 54